



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

보건학석사 학위논문

# 지리가중회귀모형을 이용한 지역별 차별 출산력 분석

Regional Fertility Differences in South Korea:  
A Geographically Weighted Regression Approach

2017년 8월

서울대학교 보건대학원  
보건학과 보건학 전공  
정 명 구

# 지리가중회귀모형을 이용한 지역별 차별 출산력 분석

Regional Fertility Differences in South Korea:  
A Geographically Weighted Regression Approach

2017년 8월

서울대학교 보건대학원  
보건학과 보건학 전공  
정 명 구

## 국문초록

대한민국은 2000년 초저출산국가(합계출산율 1.3 이하)에 진입한 이후로 장기적인 저출산 현상을 경험하고 있다. 저출산 문제에 대응하기 위한 정부의 적극적인 재정투입과 출산장려정책 시행에도 불구하고 합계출산율은 2016년 1.17을 보이며 OECD국가 중 최하위를 기록하고 있다.

본 연구는 지리가중회귀모형을 활용하여 전국 시군구 252개 지역의 차별 출산력을 분석하였으며, 지역별 출산력은 유사한 출산수준을 보이는 지역끼리 서로 밀집해 있을 것이라는 가정 하에 공간효과, 즉 공간자기상관성(Spatial Autocorrelation)을 고려하여 지역별 출산요인을 분석하였다. 이는 기존의 공간적인 동질성을 가정한 전역적 회귀모형과는 달리 공간적 이질성을 고려하여 지역별 회귀모형을 적용한 것이다. 지리가중회귀모형 적용의 타당성을 입증하기 위하여 시군구별 합계출산율의 공간자기상관성 검정을 실시하였고, 그 결과에 따라 지리가중회귀모형을 사용하여 지역별 차별 출산율에 영향을 주는 출산요인에 대해 분석하였다.

분석결과를 요약하면 다음과 같다. 첫째, 2010년과 2015년 시군구 합계출산율은 전역적(global), 국지적(local)으로 공간자기상관성을 가지고 있었다. 전역적 공간자기상관성 검정을 위해 Moran's I 지수를 활용하였으며 그 결과는 0.3( $P < 0.01$ ) 이상으로 지역별 출산율이 정적인(+) 공간자기상관성을 가지고 있음을 확인하였다. 지역별 출산율의 국지적인 공간자기상관성 패턴을 분석하기 위하여 Getis-Ord  $G_i^*$  지수를 활용하였으며, 이를 통해 통계적으로 유의하게 출산수준이 비슷한 지역끼리 군집현상을 보이는 것을 확인하였다. 높은 출산율을 가진 지역끼리 군집현상을 보이는 핫스팟(Hot-Spot) 지역은 전라남도에서 주로 일어났으며, 낮은 출

산율의 지역적 군집현상이 일어나는 콜드스팟(Cold-Spot) 지역은 서울특별시와 부산광역시, 그리고 경기도와 부산광역시의 일부 대도시에서 확인되었다. 이를 통해 지역별 출산율의 전역적, 국지적 공간자기상관성을 통계적 검정을 통해 확인하였으며, 이에 따라 지리가중모형을 이용한 지역별 출산력 분석의 타당성을 확보하였다.

둘째, 지역별 출산요인에 대한 분석을 위하여 지리가중회귀모형을 이용하였다. 연구의 종속변수는 출산력 수준을 비교하기 위해 대표적으로 사용되는 지표인 합계출산율을 사용하였으며, 독립변수는 문헌고찰과 국가통계포털에서 수집이 가능한 시군구별 인구·결혼, 지역경제와 출산환경, 여성, 노인 관련 변수를 선정하였다. 그중 지리가중회귀모형 분석을 위해 다중공선성 문제가 발생하지 않는 변수를 최종 선정하여 분석을 실시하였다. 그 결과 인구·결혼 변수에서 2010년 가임기여성 관련 변수가 출산율에 미치는 영향의 지역별 차이가 상대적으로 컸기 때문에 지역별 차별 출산율의 원인으로 작용했을 가능성이 큰 것으로 확인되었다. 하지만 2015년에는 가임기여성 관련 변수 효과의 지역별 차이가 줄어들어 차별출산력을 설명할 수 있는 변수로 선정하기 어려웠다. 여성관련 변수는 2010년, 2015년 모두 여성취업인구비율의 효과가 지역별로 차이가 컸으며, 여성취업인구비율이 높을수록 출산율이 높아지는 양(+)의 효과를 보이는 지역과 반대로 음(-)의 효과를 보이는 지역도 있었다. 지역경제변수에서는 2010년, 2015년에 걸쳐 높은 아파트 전세값이 출산율 낮추는 음(-)의 효과가 서울과 수도권지역에서 지속적으로 강하게 나타났고, 아파트 전세값이 연구 변수 중 공간자기상관성이 가장 높은 변수로 확인되어 수도권지역에 집중된 저출산 현상의 원인으로 추정할 수 있었다. 반면 노인관련 변수는 지역별 차별 출산력에 미치는 영향이 매우 적은 것으로 나타났다.

본 연구의 한계점은 첫째, 2010년과 2015년 2개년도 자료를 사용하여 분석을 실시하였기 때문에 장기적인 변화에 대한 분석을 실시할 수 없었다. 따라서 장기적인 데이터를 활용하여 시계열분석과 공간분석을 결합하여 분석한다면 지리가중모형 분석결과에 대한 설명력을 높일 수 있을 것으로 생각된다. 둘째, 지리가중회귀모형은 다중공선성의 문제가 발생하면 연산이 어려워지기 때문에 변수들 간의 다중공선성문제를 해결할 수 있는 공간분석방법을 활용한다면 모형의 적합성과 설명력이 개선된 지역별 차별 출산력 분석이 가능할 것이다.

위의 한계점에도 불구하고 본 연구는 국내 출산 연구에서 부족했던 지역별 출산력에 대한 공간자기상관성 분석과 지역별 차별 출산력에 대한 계량 지리학적(Quantitative Geography) 접근을 시도하였다는 측면에서 의의가 있다. 위에서 언급된 한계점이 보완된 출산력 연구들이 수행된다면 지역별 특성을 반영한 지방자치단체들의 효과적인 출산장려 정책 수립과 출산환경조성에 기여할 수 있을 것이다.

주요어: 합계출산율, 저출산, 지역별 차별 출산력, 지리가중회귀모형,  
GIS, 공간분석, 계량 지리학  
학 번: 2015-24081

## 목 차

|                                   |    |
|-----------------------------------|----|
| 제1장 서론 .....                      | 1  |
| 제1절 연구배경 및 필요성 .....              | 1  |
| 제2절 연구목적 .....                    | 6  |
| 제2장 문헌고찰 및 연구가설 .....             | 7  |
| 제1절 문헌고찰 .....                    | 7  |
| 1. 출산요인 선행연구 검토 .....             | 7  |
| 2. 지역별 차별 출산력 선행연구 검토 .....       | 8  |
| 3. 지리가중회귀 모형을 활용한 선행연구 검토 .....   | 12 |
| 제2절 연구가설 .....                    | 14 |
| 제3장 연구방법 .....                    | 15 |
| 제1절 연구자료 및 연구지역 .....             | 15 |
| 1. 연구자료 .....                     | 15 |
| 2. 연구지역 .....                     | 15 |
| 제2절 연구변수 .....                    | 16 |
| 1. 종속변수 .....                     | 16 |
| 2. 독립변수 .....                     | 17 |
| 제3절 분석방법 .....                    | 20 |
| 1. 연구대상지역의 일반적인 특성 및 기초통계분석 ..... | 20 |
| 2. 상관분석 .....                     | 20 |
| 3. 공간자기상관 분석 .....                | 21 |

|                                       |    |
|---------------------------------------|----|
| 제4장 연구결과 .....                        | 33 |
| 제1절 연구대상지역의 일반적 특성 및 기본통계 분석결과 .....  | 33 |
| 제2절 상관관계 분석결과 .....                   | 38 |
| 제3절 공간자기상관 분석결과 .....                 | 39 |
| 1. 지역별 출산율과 관련변수의 전역적 공간자기상관 검정 ..... | 39 |
| 2. 지역별 출산율과 관련변수의 국지적 공간자기상관 검정 ..... | 42 |
| 제4절 지리가중회귀모형 분석결과 .....               | 50 |
| 1. 연구변수 선정 .....                      | 50 |
| 2. 인구·결혼변수와 지역출산율 관계 분석 .....         | 52 |
| 3. 여성관련 변수와 지역출산율 관계 분석 .....         | 56 |
| 4. 지역경제 및 출산환경 변수와 지역출산율 관계 분석 .....  | 60 |
| 5. 노인관련 변수와 지역출산율 관계 분석 .....         | 64 |
| 제5장 결론 및 논의 .....                     | 68 |
| 제6장 연구의 한계점 .....                     | 71 |
| 참고 문헌 .....                           | 72 |
| Abstract .....                        | 77 |



## 표 목 차

|   |    |
|---|----|
| 표 1. 시도별 합계출산율 .....                          | 3  |
| 표 2. 시군구별 합계출산율 .....                         | 4  |
| 표 3. 경기도 시군 출산환경지수 지표체계 .....                 | 10 |
| 표 4. 제주특별자치도 출산영향 평가지표 .....                  | 11 |
| 표 5. 연구변수 정리 .....                            | 19 |
| 표 6. 전역적 모델, 국지적 모델 비교 .....                  | 26 |
| 표 7. 연구목적 및 연구방법 .....                        | 32 |
| 표 8. 연구변수 기초통계량(2010년도) .....                 | 36 |
| 표 9. 연구변수 기초통계량(2015년도) .....                 | 37 |
| 표 10. 독립변수와 출산율과의 상관관계(2010, 2015년도) .....    | 38 |
| 표 11. 시군구 합계출산율에 대한 Moran' s I 분석 결과 .....    | 41 |
| 표 12. 2010년 인구·결혼 변수 GWR/OLS 분석결과 .....       | 54 |
| 표 13. 2015년 인구·결혼 변수 GWR/OLS 분석결과 .....       | 54 |
| 표 14. 2010년 여성 변수 GWR/OLS 분석결과 .....          | 57 |
| 표 15. 2015년 여성 변수 GWR/OLS 분석결과 .....          | 58 |
| 표 16. 2010년 지역경제 및 출산환경 변수 GWR/OLS 분석결과 ..... | 61 |
| 표 17. 2015년 지역경제 및 출산환경 변수 GWR/OLS 분석결과 ..... | 62 |
| 표 18. 2010년 노인변수 GWR/OLS 분석결과 .....           | 65 |
| 표 19. 2015년 노인변수 GWR/OLS 분석결과 .....           | 66 |

## 그 립 목 차

|  |    |
|--|----|
| 그림 1. 연도별 전국 합계출산율 .....               | 2  |
| 그림 2. 전국 시군구 행정구역 지도 .....             | 16 |
| 그림 3. Moran' s I 값에 따른 공간자기상관성 .....   | 22 |
| 그림 4. 지리가중회귀모형의 공간가중 개념 .....          | 29 |
| 그림 5. 대역폭(Bandwidth) 설정 방법 비교 .....    | 30 |
| 그림 6. 2010년, 2015년 지역별 합계출산율 .....     | 35 |
| 그림 7. 인구변수 Getis-Ord Gi* 분석결과 .....    | 45 |
| 그림 8. 여성 관련변수 Getis-Ord Gi* 분석결과 ..... | 46 |
| 그림 9. 지역경제변수 Getis-Ord Gi* 분석결과 .....  | 47 |
| 그림 10. 노인변수 Getis-Ord Gi* 분석결과 .....   | 49 |
| 그림 11. 다중공선성 확인 및 최종 연구변수 선정 .....     | 51 |
| 그림 12. 인구·결혼 변수 회귀계수 .....             | 55 |
| 그림 13. 여성관련 변수 회귀계수 .....              | 59 |
| 그림 14. 지역경제 및 출산환경 회귀계수 .....          | 63 |
| 그림 15. 노인 변수 회귀계수 .....                | 67 |

# 제1장 서론

## 제1절 연구배경 및 필요성

대한민국의 합계출산율은 그림 1과 같이 2000년 1.3<sup>1)</sup>이하로 떨어지면서 초저출산국가에 진입한 후로, 지속적인 감소추세를 보이고 있다. 장기화 되고 있는 저출산 현상은 ‘인구절벽’이라는 생소한 용어를 한국사회에 확산시켰으며, 성인 10명중 9명은 인구절벽이 우리사회에 부정적인 영향을 줄 것이라고 인식을 할 정도로 인구감소에 대한 한국사회의 염려가 높아지고 있다. 인구감소와 가장 밀접한 관계가 있는 저출산 문제 회복을 위하여 한국정부는 저출산의 장기화에 따른 사회·경제적 파급효과에 대응하고자 2006년부터 적극적인 출산정책을 추진하고 있다(이삼식, 2016). 이미 제1차 저출산·고령사회기본계획(2006-2010)과 제2차 저출산·고령사회기본계획(2011-2015)이 추진되었고, 2016년부터 제3차 저출산·고령사회기본계획(2016-2020)을 시행하고 있다. 이러한 정부의 출산정책과 2020년까지 합계출산율 1.5를 목표로 하는 3차 기본계획의 수립에도 불구하고, 2016년 전국 합계출산율은 1.17을 기록하여 정부의 목표달성을 낙관적으로만 예상하기 어려운 실정이다.

---

1) 합계출산율은 여성 1명이 평생 동안 낳을 것으로 예상되는 평균 출생아 수를 나타낸 지표이다. 따라서 합계출산율 1.3의 의미는 여성 1명이 일생동안 1.3명의 아이를 출산할 것으로 예상한다는 의미이다.

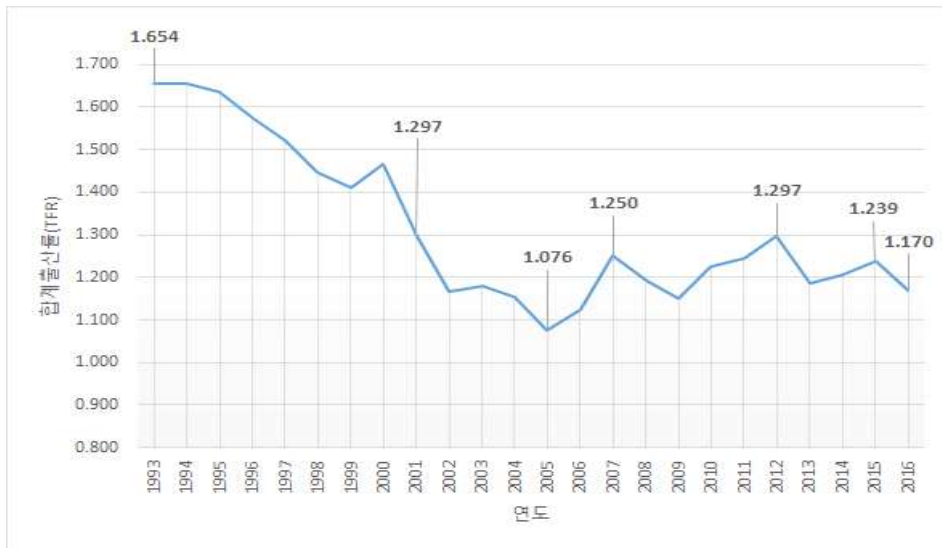


그림 1. 연도별 전국 합계출산율

효과적인 출산정책 수립을 위해서는 전국 출산율을 결정하는 요인을 분석하여 활용하는 것도 중요하지만, 광역자치단체나 기초자치단체 수준의 지역별 특성을 고려한 지역별 출산요인에 대한 분석 또한 중요하다 (이헌창 외, 2012). 보건복지부 역시 지방자치단체의 차별적인 출산정책의 중요성을 인식하여 매년 지방자치단체 출산장려정책사례집(구 인구정책사례집)을 발간하고 있다. 우리나라의 출산율 양상의 특징 중 하나는 대도시를 중심으로 상대적으로 낮은 출산율이 집중되는 지역별 차별 출산율 현상이 고착되는 것이다. 표 1은 2005년, 2010년, 2015년 시도별 출산율을 나타내며 대도시인 서울을 비롯한 광역시는 낮은 출산율을 보이는 반면, 도 단위의 행정구역에서는 상대적으로 높은 출산율을 보이는 특징을 확인할 수 있다.

표 1. 시도별 합계출산율

| 구분 | 2005년 |       | 2010년 |       | 2015년 |       |
|----|-------|-------|-------|-------|-------|-------|
|    | 지역    | TFR   | 지역    | TFR   | 지역    | TFR   |
| 1  | 제주도   | 1.296 | 전라남도  | 1.537 | 전라남도  | 1.549 |
| 2  | 전라남도  | 1.282 | 충청남도  | 1.479 | 울산광역시 | 1.486 |
| 3  | 충청남도  | 1.256 | 제주도   | 1.463 | 충청남도  | 1.480 |
| 4  | 충청북도  | 1.187 | 경상남도  | 1.413 | 제주도   | 1.477 |
| 5  | 경상남도  | 1.182 | 충청북도  | 1.402 | 경상북도  | 1.464 |
| 6  | 울산광역시 | 1.181 | 경상북도  | 1.377 | 경상남도  | 1.437 |
| 7  | 강원도   | 1.177 | 전라북도  | 1.374 | 충청북도  | 1.414 |
| 8  | 전라북도  | 1.175 | 울산광역시 | 1.369 | 전라북도  | 1.352 |
| 9  | 경기도   | 1.173 | 강원도   | 1.313 | 강원도   | 1.311 |
| 10 | 경상북도  | 1.166 | 경기도   | 1.309 | 대전광역시 | 1.277 |
| 11 | 대전광역시 | 1.100 | 광주광역시 | 1.223 | 경기도   | 1.272 |
| 12 | 광주광역시 | 1.097 | 인천광역시 | 1.214 | 인천광역시 | 1.216 |
| 13 | 인천광역시 | 1.065 | 대전광역시 | 1.206 | 대구광역시 | 1.216 |
| 14 | 대구광역시 | 0.995 | 대구광역시 | 1.109 | 광주광역시 | 1.207 |
| 15 | 서울특별시 | 0.922 | 부산광역시 | 1.045 | 부산광역시 | 1.139 |
| 16 | 부산광역시 | 0.879 | 서울특별시 | 1.015 | 서울특별시 | 1.001 |

※ 음영: 서울특별시와 광역시 지역, 출처: 통계청자료를 활용하여 구성

표 2. 시군구별 합계출산율

| 구분  | 2005년     |       | 2010년     |       | 2015년     |       |
|---|-----------|-------|-----------|-------|-----------|-------|
|   | 지역        | TFR   | 지역        | TFR   | 지역        | TFR   |
| 상위<br>5개 지역                               | 강원<br>화천군 | 1.710 | 전북<br>진안군 | 2.41  | 전남<br>해남군 | 2.464 |
|   | 강원<br>인제군 | 1.598 | 전남<br>강진군 | 2.339 | 강원<br>인제군 | 2.161 |
|   | 전남<br>영암군 | 1.591 | 전남<br>영암군 | 2.172 | 전남<br>영암군 | 2.107 |
|   | 강원<br>양구군 | 1.583 | 강원<br>인제군 | 2.004 | 전남<br>장성군 | 2.101 |
|   | 울산<br>북구  | 1.534 | 전남<br>보성군 | 2.003 | 부산<br>강서구 | 1.982 |
|   | 평균        | 1.603 | 평균        | 2.186 | 평균        | 2.163 |
| 하위<br>5개 지역                               | 부산<br>영도구 | 0.727 | 서울<br>종로구 | 0.837 | 부산<br>서구  | 0.899 |
|   | 광주<br>동구  | 0.724 | 대구<br>서구  | 0.828 | 부산<br>영도구 | 0.899 |
|   | 대구<br>중구  | 0.722 | 대구<br>남구  | 0.822 | 서울<br>강남구 | 0.857 |
|   | 서울<br>강남구 | 0.712 | 경북<br>울릉군 | 0.796 | 서울<br>관악구 | 0.833 |
|   | 부산<br>중구  | 0.696 | 부산<br>서구  | 0.780 | 서울<br>종로구 | 0.813 |
|   | 평균        | 0.716 | 평균        | 0.813 | 평균        | 0.860 |
| [상위5개 지역 평균 TFR]<br>-<br>[하위5개 지역 평균 TFR] |           | 0.887 |           | 1.373 |           | 1.303 |

※ 출처: 통계청자료를 활용하여 구성

또한 지역별 차별 출산율의 추이를 시군구 단위로 살펴보면 지역별 출산율 차이가 확대되는 것을 확인할 수 있다. 표 2는 2005년, 2010년, 2015년의 시군구별 합계출산율 중 출산율이 높은 상위 5개 지역과 출산율이 낮은 하위 5개 지역을 통계청 자료를 활용하여 정리한 것이다. 상

위 5개 지역의 합계 출산율 평균과 하위 5개 지역의 합계출산율 평균을 계산하여 그 차이를 확인하면 2005년 0.887에서 2015년 1.303으로 확대되는 것을 확인할 수 있다. 시군구의 출산율의 차이 확대가 전국 합계출산율의 하락현상과 대조적이라는 점에서 효과적인 출산정책 수립을 위해 지역별 차별 출산율에 대한 분석이 필수적이라고 할 수 있다. 또한 출산율이 높은 시군구 지역은 전라도와 강원도에서 지속적으로 확인되고, 출산율이 낮은 시군구 지역은 지속적으로 서울특별시와 부산광역시 그리고 대구광역시에서 확인되는 것을 보았을 때, 출산율이 지리적인 혹은 공간적 상관성이 있을 것으로 생각할 수 있다. 다시 말해, 지속적으로 출산수준이 유사한 지역끼리 밀집해 있는 현상을 확인할 수 있으며, 이러한 현상을 지리학에서는 공간자기상관성(Spatial Autocorrelation)을 갖는다고 한다. 공간적 자기상관이 일어나는 이유 중 하나는 공간의 이질성으로, 지역 간의 특성(예: 도시, 농촌)으로 인해 지역적 차이가 발생하는 것을 말한다. 예를 들어, 출산율에 대한 동일한 정책 효과가 모든 지역에서 동일하게 일어나지 않고, 동일한 정도로 변화하더라도 그 영향이 지역에 따라 달라진다는 것이다. 또한 이러한 공간자기상관성은 지역 간의 파급효과(Spillover Effect)를 만들어 인접한 지역끼리 유사한 형태의 현상이 일어나는 확산효과를 보이기도 한다.

공간적 이질성으로 인한 공간자기상관성 문제를 해결하기 위하여 계량지리학 연구에서는 변수간의 관계를 추정하는 회귀계수가 지역 간에 서로 다르다는 것을 전제로 지역별 회귀모델을 도출하여 분석을 시도하고 있다. 따라서 본 연구는 도시를 중심으로 낮은 출산율이 일어나는 공간자기상관성을 고려한 지리가중회귀모형을 이용하여 지역별 출산력 분석을 시도하고자 한다.

## 제2절 연구목적

본 연구는 앞서 언급한 것과 같이, 공간자기상관성을 고려하여 지역별 생산력의 차이를 분석하되 생산력 수준을 나타내는 가장 대표적인 지표인 지역별 합계출산율을 종속변수로 설정하고, 이와 관련 있는 출산요인을 찾아 효과적인 지역단위의 출산정책 방향을 제시하는데 목적이 있다. 이를 위해 우리나라 전체 시군구 단위로 수집된 출산관련 자료를 GIS(Geographic Information System)을 사용하여 시각화 할 것이며, 지역별 차별 출산율의 분석을 위해 계량 지리학(Quantitative Geographical) 연구방법의 하나인 지리가중회귀모형(Geographically Weighted Regression Model)을 활용할 것이다. 따라서 본 연구의 구체적인 목적은 다음과 같다.

첫째, 시군구 지역별 차별 출산율을 확인하고 공간적 출산율 패턴을 확인한다.

둘째, 지역별 출산율과 관련변수들에 대한 공간자기상관성을 확인하고, 이로 인한 지역별 출산수준의 군집현상이 일어나는지를 확인한다.

셋째, 지역별 차별 출산력에 영향을 미치는 요인을 지리가중회귀모형을 활용하여 분석한다.



## 제2장 문헌고찰 및 연구가설

### 제1절 문헌고찰

#### 1. 출산요인 선행연구 검토

저출산 문제가 심화되면서 이에 대한 국내 연구가 활발히 진행되고 있다. 선행 연구를 살펴보면, 출산요인으로 인구학적 요인인 초혼연령과 혼인율 및 교육수준, 경제학적 요인인 출산의 효용과 비용 관점의 분석, 사회문화적 요인인 결혼·출산에 대한 가치관 변화 등의 효과를 강조하고 있으며, 출산 문제는 다양한 요인들이 복합적으로 작용하여 나타나는 결과임을 확인할 수 있다. 박병일(2004)은 1970~2002년의 인구변동과 우리나라의 합계출산율에 영향을 미치는 변수들 간의 관계를 분석하여 합계출산율과 인구밀도 간에 역의 상관관계가 있음을 확인하였다. 혼인율 하락과 초혼연령 상승의 경제효과를 연구한 오창섭·최성혁(2012)은 고용불안정성과 소득의 불균형이 젊은 남녀의 결혼을 제약하는 것으로 보고하였으며, 주택비용 증가와 고용불안은 남성의 결혼참여를 어렵게 하는 것으로 나타났다. 성낙일·박선권(2012)은 자녀양육 환경이 출산율에 미치는 영향을 분석해서 보육시설의 양적 규모가 출산율에 긍정적 영향을 미친다고 보고하였다. 이삼식·최효진(2012)의 연구에서는 주거형태별 출산행태를 분석하였는데, 자가에 비해 전세인 경우, 단독주택보다 아파트인 경우, 출생아 수 증가에 부정적인 효과가 있는 것으로 보고되었고, 주거면적이 아주 넓은 경우 출생아 수를 증가시키는 효과가 있는 것으로 나타났다. 또한 부모의 교육수준에 따른 여성의 초산시기와 출산간격이 출산율에 영향을 미친다는 연구결과도 있었다(류기철·박영화, 2009; 신

혜원 외, 2012; 우해봉, 2012). 김민영, 황진영(2016)은 2009~2013년 우리나라 16개 시도의 주택가격 자료를 활용해서 주택가격이 합계출산율에 음의 영향을 미치는 것으로 보고하였다. 위의 연구들은 대부분 개인단위로 연구를 실시하였기 때문에 지역단위의 출산요인을 분석하는데 한계점을 가지고 있다. 따라서 지역단위의 출산력 차이에 대한 선행연구를 검토하도록 하겠다.

## 2. 지역별 차별 출산력 선행연구 검토

출산력은 사회경제적 특성에 따라 구분되는 변수이므로 집단별, 지역별로 다르다. 이와 같은 다양한 기준에 따른 출산력의 차이를 인구학에서는 ‘차별 출산력’으로 개념화하고 있다. 출산요인의 인과구조를 파악하기 위해서는 특정한 인구 사회적 특성을 가진 집단이 특정 지역에서 어떠한 출산 행태를 보이는지에 대한 지역별 차별 출산에 대한 연구가 필요하다(최은영·박은실, 2010). 특히 출산의 문제가 개인의 문제가 아닌 사회적인 문제로 인식되면서, 개인 및 가구의 특성보다는 지역단위의 특성을 분석하는 연구가 활발해지고 있다. Kulu and Vikat(2007)는 인구수가 많은 지역이 적은 지역보다 낮은 출산율을 보이며, 출산율의 감소속도 역시 빨라진다고 하였다. 또한 Hank(2001), Boyle(2003)는 사회·경제·문화적 요인들의 상호작용에 의해 소도시가 대도시보다 높은 출산율을 나타낸다고 하였다.

지역별 차별출산력에 대한 국내연구를 살펴보면, 은기수(1997)는 1990년도까지 합계출산율이 인구대체수준(TFR 2.1) 이하로 떨어졌지만 도시와 농촌의 차별출산력이 존재하기 때문에 군·구의 출산력이 시·도보다 높게 나타나고 있음을 분석하였다. 시도별로는 전라도 지역이 높은 출산율을 유지하고, 서울과 부산이 매우 낮은 출산율을 보이고 있음을 확인

하였다. 이시원 외(2006)는 1995-2002까지의 시·군 지방자치단체를 대상으로 출산감소 요인을 분석하였다. 그 결과, 시 지역의 경우 모(母)의 취업률이 높고, 첫 출산연령이 높으며, 이혼건수가 많은 지방자치단체일수록 출산율 감소가 큰 반면, 군 지역은 복지비 지출이 높고, 모자보건사업이 많이 시행되는 지방자치단체의 출산율 감소가 적은 것으로 분석하였다. 유삼현(2007)은 전국 시군구 234개 지역의 차별출산력을 살펴보기 위해 사회경제적 수준과 제2차 출산력 변천의 주요 결정요인인 노동시장의 변화, 가족의 형성과 해체 등을 지역 단위로 측정하여 합계출산율과의 관계를 분석하였다. 분석결과 사회경제적 수준이 높을수록 합계출산율이 감소하였고, 가족형성이 제한되거나 해체되는 경향이 높은 지역일수록 합계출산율도 함께 낮아졌다. 또한 노동시장의 불안정성이 높은 지역일수록 합계출산율이 낮은 것으로 분석되었다. 최은영, 박영실(2010)의 연구에 따르면, 시군구별 합계출산율은 군 지역이 시 지역보다 합계출산율이 높았으며, 휴전선 인근지역과 남해안의 산업도시에서 상대적으로 높은 합계출산율이 나타난다고 하였다. 또한 합계출산율은 국제결혼의 비율, 일자리 증가율과 정적(+)상관관계를 갖고, 미혼율은 부적(-)상관관계를 보였다.

지역별 차별 출산력과 관련하여 지역의 출산환경을 측정할 수 있는 지표 및 지수개발 역시 연구되고 있으며, 이를 통해 지역별 출산행위에 영향을 미치는 다양한 요인들을 파악하여 지역 특성을 고려한 출산정책 수립을 위한 자료로 활용되고 있다. 고지영 외(2012)는 경기도가 31개의 행정구역으로 나누어진 지역으로서, 자치단체별 출산 수준의 편차가 매우 큰 것에 주목하여, 경기도 지역 출산에 영향을 미치는 출산환경을 종합적으로 분석하기 위해 표 3과 같은 출산환경지수 지표를 개발하였다.

표 3. 경기도 시군 출산환경지수 지표체계

| 구분       | 지표                       |
|----------|--------------------------|
| 인구 결혼    | 청년인구 비율                  |
|          | 여성 초혼연령                  |
|          | 조혼인율                     |
| 가치관      | 다자녀출산 가치관                |
| 경제적 여건   | 청년고용률                    |
|          | 지역소득수준                   |
|          | 지역 내 전세가격                |
|          | 1인당 월평균 교육비              |
| 일 가정 양립  | 기혼여성의 경제활동 참가율           |
|          | 장시간 일하는 취업자 비율           |
|          | 장시간 통근자 비율               |
|          | 가사노동시간 성비                |
| 지역양육 인프라 | 지역 내 보육시설 제공수준           |
|          | 지역 내 교육기관 제공수준           |
|          | 지역 내 가족지원기관 제공수준         |
|          | 지역 내 가족, 보육, 청소년 분야 예산비중 |

정영태 외(2012)는 제주지역의 저출산 현상 극복을 위해 생애단계별 출산 장애 요인을 파악하고, 이를 대응할 수 있도록 표 4와 같이 출산영향 평가지표를 개발 하였다.

표 4. 제주특별자치도 출산영향 평가지표

| 구분              | 항목             | 지표                                 |
|-----------------|----------------|------------------------------------|
| 출산<br>환경        | 직장문화 및<br>근무환경 | 여성 취업 장애요인에 대한 태도<br>(육아 및 가사부담)   |
|                 | 의료와 보건환경       | 15~49세 여성인구 10만 명당 분만실보유<br>산부인과 수 |
|                 | 출산장려정책         | 가임여성 인구당 '저출산 시책 예산' 규모            |
| 양육<br>환경        | 보육환경           | 0~4세 인구당 일반 보육시설(어린이집) 수           |
|                 | 교육환경           | 시도별 월평균 사교육비 지출액                   |
|                 | 주거환경           | 시도별 국민주택 규모(85㎡) 이상 주택 수 비율        |
|                 | 안전             | 범죄위험에 대한 안전 인식                     |
| 개인적<br>영향<br>요인 | 가치관            | 결혼에 대한 태도                          |
|                 | 소득수준           | 현 소득에 대한 만족도                       |
|                 | 건강수준           | 여성의 주관적 건강수준에 대한 만족도               |
| 인구학적 요인         |                | 총 여성인구대비 15~44세 가임여성 인구 비율         |

앞서 살펴본 선행연구들은 지역별 차별출산력 문제에 주목하고 개인을 넘어 지역을 단위로 하여 출산요인을 규명하고자 했다는 점에서 의의를 갖는다. 다만 지역이라는 의미에서 중요한 ‘공간’ 요인을 고려하고 있지 못하고 있으며, 개별 관찰치가 주어진 공간상에 무작위적(random)이고 독립적으로 분포한다는 가정 하에 분석이 이루어졌다. Tobler(1970)의 지리학 1법칙인 ‘모든 것은 다른 모든 것과 관련을 맺지만, 가까운 것이 먼 것보다 더 밀접한 관련을 갖는다’는 공간이나 특정지역에 위치

한 대상은 공간에 무작위적이고 독립적으로 분포하지 않고 서로 영향을 주고받으며, 그 영향은 공간상의 거리나 인접성이 높을수록 더 커지는 경향이 많다. 따라서 사회경제현상과 인구현상을 분석할 때, 공간상에서 나타나는 특성인 공간상호작용, 즉 공간자기상관(Spatial Autocorrelation)을 고려해야하며, 지역적 차별 출산력에도 이러한 공간요인을 반영한 연구가 필요하다.

### 3. 지리가중회귀 모형을 활용한 선행연구 검토

지리가중회귀모형(GWR: Geographically Weighted Regression Model)은 그 결과를 GIS(Geographic Information System) 지도로 제시함으로써 데이터의 공간적 상호작용을 시각적으로 확인할 수 있다는 장점을 가지고 있다(Downey L., 2006; 조동기, 2009). 국내에서는 지리가중회귀모형이 지리적 효과와 공간자기상관이 높은 부동산 혹은 주택가격분야에 적용이 되고 있다. 이와 관련하여 부산시, 서울시, 강남구, 농촌지역 등 특정 지역의 주택·아파트 가격의 변동요인을 GWR 모형을 이용하여 분석하고 있으며, 기존의 일반회귀모형보다 결정계수 등을 비교했을 때 GWR 분석의 설명력과 적합성이 높다는 연구결과들이 있다(오윤경 외, 2014; 강창덕, 2010; 정건섭, 2011; 박헌수, 2011; 송용철·박헌수, 2012). 국외에서도 Faver and Yeates(2006)는 GWR 모형이 OLS(Ordinary Least Squares) 보다 통계적으로 설명력이 높으며, 공간적 특성과 이질성을 파악하는데 더욱 효과적이라고 하였다. 한편 김혜영·전철민(2012)은 지가 분석을 위해 GWR 모형과 OLS 모형을 비교 분석한 결과, 공간효과에 대한 존재를 확인하였으나 OLS와 GWR의 분석방법과 결과에 대한 특징이 다르기 때문에 두 모델을 병행하여 분석하는 것이 좋다고 하였다.

지리가중회귀모형은 보건·의료 분야에서도 활용되고 있다. 국외에서

는 Wen et al.(2010)이 가난과 비만의 연관성에 대한 지역별 변이 연구를 시행하였으며, Lin and Wen(2011)은 텅기열의 지역별 발생률의 차이를 파악하기 위해 GWR 모형을 활용하였다. 국내에서는 박일수 외(2013)가 지역별 심정지 발생요인 분석을 위해 지리가중회귀모형을 전통적인 OLS 모형과 비교하여 모형의 우수성을 확인하였으며, 심정지 발생 예방관리를 위한 우선 사업 순위를 선정하였다. 김유미·강성홍(2014)은 지리가중회귀모형을 통해 전국 시군구별 지역사회건강조사 자료를 활용하여 지역별 비만을 변이에 관한 연구를 실시하였고, 결과를 바탕으로 지역별 비만관리 사업의 우선순위를 제안하였다.

최근에는 국내 인구와 관련된 부분에서도 지리가중회귀모형이 활용되고 있다, 최돈정·서용철(2012)은 장수인구분포의 공간적 변이를 GWR 모형을 활용하여 지역별 장수의 환경생태학적 요인을 분석하였다. 윤정미·최돈정(2015)은 서울시 유동인구 분포의 공간 패턴과 토지이용을 분석하기 위해 GWR을 활용하였으며, 그 결과 서울시 유동인구 분포가 토지이용에 변화를 주는 것으로 확인하였다. 김리영·서원석(2016)은 학령별 인구의 공간적 분포와 지역특성을 지리가중회귀모형을 통해 분석하였으며 지역별 교육밀도와 지역특성을 고려한 교육수요를 추정해야 한다는 결론을 도출하였다.

지리가중회귀모형을 활용한 출산력 분석은 국내에서는 부족한 실정이나 국외에서는 시도 되고 있다. Oguz and Melih(2006)는 지리가중회귀모형을 이용하여 터키의 출산수준을 분석하였으며, 터키의 출산감소현상이 전역에 일어나는 것이 아니라, 공간적 자기상관을 기반으로 한 지역별 차별출산수준이 존재한다고 하였다. Kamata and Iwasawa(2010)는 GWR 모형을 이용하여 2005년과 2010년의 일본의 출산율 추세를 분석하였으며, 출산율 추세는 전국에서 동일한 패턴으로 나오기 보다는 지역의 역

사적, 지리적 배경에 따라 다르게 나타날 수 있음을 확인하였다.

따라서 지리가중회귀모형은 지역별 특수성을 분석하는데 적합한 모형이며, 다양한 분야에서 활용되고 있다. 국외에서는 이미 지역별 차별 출산력 분석을 GWR 모형을 통해 잔행 하고 있으나, 국내에서는 공간분석을 통한 출산력 분석이 부족하다. 본 연구에서는 GWR을 활용한 국내 시군구 지역 차별 출산력에 대한 분석을 시도하고자 한다.

## 제2절 연구가설

본 연구는 지역별 출산율이 공간적으로 상호 관련이 있다고 가정하였으며, 이를 검증하기 위해 전국 시군구 자료를 활용하여 아래와 같은 연구가설을 설정하였다.

연구가설 1: 지역별 출산율은 공간적으로 자기상관성을 가지고 있다.

연구가설 2: 출산율이 유사한 지역끼리 군집현상을 보인다.

연구가설 3: 지역별 인구·결혼, 여성, 지역경제의 특성은 지역별 차별 출산율과 유의미한 연관성을 갖는다.



## 제3장 연구방법

### 제1절 연구자료 및 연구지역

#### 1. 연구자료

본 연구에서는 통계청에서 운영하는 국가통계포털(kosis.kr) 사이트에서 제공된 전국 시군구 통계지표 중 2010년과 2015년 데이터를 수집하여 분석을 실시하였다. 수집된 지표 설명과 변수에 대해서는 다음 절에 자세히 설명 하도록 하겠다.

수집된 데이터를 지도에 구현하기 위하여 통계청 통계지리정보서비스(sgis.kostat.go.kr)에서 제공되는 2015년도 센서스용 행정구역경계(시군구)의 지도 데이터(shp. 파일)를 활용하였으며, 이를 수집된 데이터와 결합하여 GIS 분석을 위한 데이터 셋을 구축하였다.

#### 2. 연구지역

본 연구의 대상 지역은 대한민국 시군구 행정구역이며, 그림 2에 나뉘어져 있는 총 252개의 행정구역을 대상으로 분석을 실시하였다.

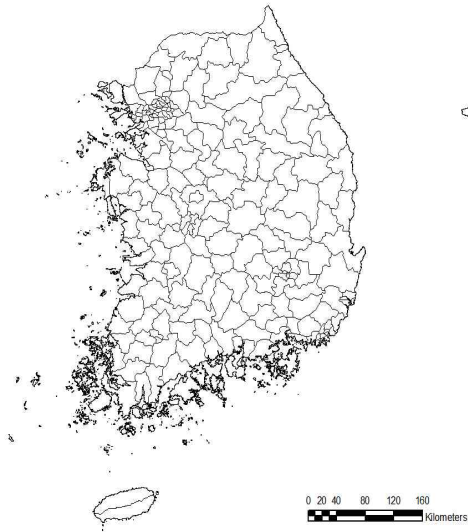


그림 2. 전국 시군구 행정구역 지도

## 제2절 연구변수

본 연구의 연구변수는 문헌고찰을 통해 확인된 지역별 차별출산력과 출산요인 중 국가통계포털 사이트에서 2010년, 2015년 시군구별 자료수집이 가능한 변수를 고려하여 선정하였다. 수집 가능한 시군구별 데이터가 제한적이었으나 문헌고찰을 통해서 확인된 주요 출산요인인 인구·결혼, 여성, 지역경제·출산환경 등에 대한 주요 변수를 수집하여 분석을 실시하였다. 연구에 활용된 변수는 다음과 같다.

### 1. 종속변수

종속변수는 시군구별 합계출산율(TFR: Total Fertility Rate)이며, 이는 여성 1명이 평생 동안 낳을 것으로 예상되는 평균 출생아 수를 나타낸 지표이다. 이는 출산력 수준을 비교하기 위하여 대표적으로 활용되는 지표로서 일반적으로 연령별 출산율(ASFR: Age-Specific Fertility Rate)의

합으로 계산된다. 합계출산율의 수치는 합계출산율이 높을수록 한 여성이 평생 동안 출산할 것으로 예상되는 자녀수가 많다는 것을 의미한다.

## 2. 독립변수

독립변수는 4개의 주제(인구·결혼, 여성, 지역경제·출산환경, 노인)와 출산율과의 관계를 살펴볼 것이다. 첫 번째 주제인 인구·결혼과 출산율과의 관계를 분석하기 위하여 7가지 변수를 선택하였다. 첫 번째는 조혼인율로서 ‘조혼인율 = (연간 총 혼인건수 ÷ 당해년도 주민등록연앙인구) × 1,000’의 공식으로 산출되는데, 연간 발생한 총 혼인건수를 당해년도 주민등록연앙인구(7월 1일)로 나누어 천분비로 나타낸 수치이다. 즉 조혼인율은 인구 1,000명당 혼인건수를 의미한다. 두 번째 변수는 인구밀도로 해당지역의 전체인구를 지역의 면적으로 나누어 값을 추출하였다. 세 번째 변수는 인구증가율로 전년도 인구수를 기준으로 당해년도 인구에서 전년도 인구비율을 뺀 값에서 100을 곱한 값으로 도출된다. 당해년도 인구가 전년도 인구보다 줄었을 경우 인구 증가율 값은 음(-)의 값을 갖게 된다. 네 번째 변수는 지역별 여성의 평균 초혼연령으로 처음 결혼하는 연령을 의미한다. 지역의 평균초혼연령이 높다는 것은 해당 지역 여성의 결혼시기가 상대적으로 늦다는 것을 의미한다. 다섯 번째 변수는 가임기 여성비율로서 이는 전체 여성 중 15세-49세의 여성비율을 나타내는 것으로, 실제 출산이 가능한 여성인구 비율을 말한다. 여섯 번째 가임기 여성 미혼율로서 가임기 여성 중 미혼인 여성 비율을 의미한다. 한국과 같이 결혼과 출산이 밀접한 관계를 갖는 문화의 경우, 가임기 여성 중 미혼여성은 출산을 경험할 경우가 적다고 볼 수 있다. 인구·결혼과 관련된 마지막 변수는 외국인과의 혼인비율로서 연간 전체 혼인 건수 중 한국 남편과 외국인 여성의 혼인건수를 나누어 값을 구하였다.

두 번째 주제인 여성과 관련하여 먼저 여성취업인구비율을 살펴볼 것

이다. 여성이 취업상태에 있을 경우, 일과 가정의 양립이 어렵고, 시간에 대한 기회비용으로 인해 육아와 출산에 부정적인 영향을 줄 수 있다. 두 번째는 대졸 이상의 여성 인구비율로 가임기 여성인구 중 출산을 경험할 확률이 상대적으로 낮은 15-20세의 여성인구를 빼 21-49세의 여성인구를 대상으로 값을 구하였다. 본 지표가 높다는 것은 해당지역 여성의 교육수준이 높거나 모(母)의 교육수준이 높은 지역으로 예상할 수 있다

세 번째 주제는 지역경제로서 거주지역의 경제상황은 가족의 경제활동에 영향을 주어 출산과 양육활동에 부담을 줄 수도 있다. 이를 위해 신혼부부가 우선적으로 고려하게 되는 전세가격을 살펴볼 것이다. 본 연구에서 살펴볼 전세가격은 아파트  $m^2$ 당 평균 전세가격(단위: 만원)을 나타내는 지표로서, 높은 값은 높은 전세가격이 형성된 지역으로 예상할 수 있다. 또한 인구 1000명당 사업체 수를 살펴볼 것이며, 1000명당 사업체가 많은 지역은 경제활동을 지속할 수 있는 기회를 주는 곳으로 해석할 수 있다. 지역의 출산환경을 살펴보기 위하여 유아 $\pi$ 1000명당 보육시설수를 살펴보도록 하겠다.

네 번째 주제는 노인과 관련된 변수로서, 노인인구 비율이 높은 지역은 경제활동인구와 출산이 가능한 인구가 적다는 것을 의미하며 지역의 정책도 노인에 중점을 둔 지원과 정책이 이루어질 가능성이 높다. 이를 위해 고령인구 비율을 살펴보고, 실제 고령인구비율이 출산율에 미치는 영향이 있는지 실증 분석을 실시할 것이다. 또한 노인(60세 이상) $\pi$ 1000명당 노인복지시설수를 살펴보고 노인인구 비율에 따른 복지시설수가 관련이 있는지, 그리고 이 변수가 출산율에도 영향을 미치는지 살펴볼 것이다. 이와 함께 지역 건강보험료 총액 대비 급여비용 비율을 살펴보고 지역의 높은 의료이용에 따른 출산율 관계를 살펴볼 것이다. 일반적으로 급여비용이 보험료보다 높다는 것은 지역의 경제활동인구가 적어 건강보험료 총액이 적고 보험료를 적게 내는 노인인구나 어린아이의 의료이용이 높다는 것을 의미한다.

표 5. 연구변수 정리

| 지표명      |                           | 지표정의                                  | 출처                                       |               |
|----------|---------------------------|---------------------------------------|--|---------------|
| 종속<br>변수 | 합계출산율                     | 여성 1인이 가임기간(15-49세) 동안 낳을 수 있는 평균 자녀수 | 통계청<br>인구동향조사                            |               |
|          |                           |                                       |  |               |
| 독립<br>변수 | 인구<br>·<br>결혼             | 조혼인율                                  | 혼인건수 ÷ 주민등록연앙인구<br>× 1,000               | 통계청<br>인구동향조사 |
|          |                           | 인구밀도                                  | 인구 수 ÷ 국토면적                              |               |
|          |                           | 인구증가율                                 | (당해년도 인구 - 전년도 인구)<br>÷ 전년도 인구 × 100     |               |
|          |                           | 여성 초혼연령                               | 지역별 여성 평균 초혼연령                           |               |
|          |                           | 가임기여성<br>비율                           | 가임여성(15~49세 여성)<br>÷ 전체 여성인구 X 100       |               |
|          |                           | 가임기여성<br>미혼율                          | 가임기여성미혼인구<br>÷ 전체가임기여성 X 100             |               |
|          |                           | 외국인과의<br>혼인비율                         | 한국인남편&외국인아내혼인건수<br>÷ 전체혼인건수              |               |
|          | 여성                        | 여성<br>취업인구비율                          | 15세이상여성취업인구<br>÷ 전체여성인구                  | 통계청<br>인구동향조사 |
|          |                           | 대졸 이상<br>여성인구비율                       | 만21~49세 4년제 대졸이상<br>여성인구 ÷ 만21세~49세 여성인구 |               |
|          | 지역<br>경제<br>·<br>출산<br>환경 | 아파트<br>전세가격                           | m <sup>2</sup> 당 평균 전세가격                 | 한국감정원         |
|          |                           | 인구 1000명<br>당 사업체 수                   | 사업체 수 ÷ 전체인구1000명                        | 통계청           |
|          |                           | 유아 1,000명<br>당 보육시설 수                 | 보육시설 수<br>÷ 0~4세 주민등록인구 × 1000           | 보건복지부<br>보육통계 |
|          | 노인                        | 고령인구비율                                | (65세 이상 인구<br>÷ 전체인구) × 100              | 통계청<br>인구동향조사 |
|          |                           | 노인(60세이상)<br>천명 당<br>노인복지시설 수         | 노인복지시설 수<br>÷ 60세이상인구1000명               |               |
|          |                           | 건강보험료대비<br>급여비율                       | 총 급여비용 ÷ 총 건강보험료                         | 국민건강<br>보험공단  |

## 제3절 분석 방법

본 연구의 분석방법은 다음과 같다. 첫째, 문헌고찰을 통해 선정된 변수들에 대한 일반적 특성과 기초 통계 분석을 실시하고, 둘째, 종속변수인 합계출산율과 변수들 간의 상관관계와 다중 공선성을 검토하여 지리가중회귀분석을 위한 최종 독립변수를 선정할 것이다. 셋째, 공간적 자기상관성 검정을 통해 지리가중회귀분석의 타당성을 통계적으로 입증한 후, 지리가중회귀 모형분석을 실시할 것이다. 구체적인 분석 방법은 다음과 같다.

### 1. 연구대상지역의 일반적인 특성 및 기초통계분석

문헌고찰을 통해서 선정된 지역별 차별 출산율과 관련된 지표 중 시·군·구별 자료수집이 가능한 변수를 선정하여 이에 대한 최소값, 중앙값, 평균값, 최대값을 통해 각 지표의 지역별 분포를 파악하였다. 또한 각 값을 GIS 지도와 결합하여 변수의 지역별 분포를 시각화 하였다. 기초통계에 필요한 분석은 R 3.3.1을 사용하였으며, 지표의 지역별 분포 시각화를 위해 ArcGIS 10.4.1을 사용하였다.

### 2. 상관분석

앞서 언급된 변수들과 종속변수인 합계출산율 간의 연관성을 확인하기 위해 상관분석을 실시하였다. 통계적 유의성이 확보된 변수 중 변수 간의 다중공선성이 일어나지 않는 변수를 선별하여 지리가중회귀모형의 변수로 선정하였다. 이는 변수간에 다중공선성이 높을 경우 지리가중회귀 연산에 오류가 생기기 때문이다. 따라서 ArcGIS를 제공하고 있는

ESRI(Environmental Systems Research Institute)에서 권고하는 방법에 따라 다중공선성 조건지수 (Multicollinearity Condition Number)가 30<sup>2)</sup>이하인 변수를 조합하여 최종 지리가중회귀모형 변수로 선정하였다. 상관분석과 다중공선성 확인을 위해 R 3.3.1와 ArcGIS 10.4.1을 사용하였다.

### 3. 공간자기상관 분석

앞서 살펴본 바와 같이 한국의 지역별 출산력 차이는 확대되고 있고, 대도시를 중심으로 상대적으로 낮은 출산율, 그 외의 지역에서는 상대적으로 높은 출산율을 보이는 경향이 있다. 이러한 지역별 차별출산력이 공간적으로 상관이 있는지, 즉 출산수준이 유사한 지역끼리 서로 이웃하고 있는지 확인하기 위해서 공간자기상관성 분석이 필요하다. 공간자기상관성 분석이란, 어떤 현상에 대한 공간적인 패턴을 설명하기 위해 그 현상이 일어나는 공간적 패턴이 우연인 것인지 아니면 특정한 체계에 따라서 유사한 현상을 보이는 지역끼리 군집을 이루는지 분석하는 방법이다(Fotheringham et al., 2002). 이를 위해 두 가지 분석을 실시할 것이다. 첫째로, 지역별 차별출산율의 전역적(global) 공간자기상관을 측정하기 위해 Moran' s Index를 사용할 것이다. 둘째로, 지역별 출산율이 유사한 지역끼리 국지적(local) 군집현상을 이루고 있는지 알아보기 위해 Getis-Ord Gi\* 방법을 활용할 것이다. 자세한 분석방법은 다음과 같다.

---

2) Results are unstable in the presence of local collinearity as indicated by a condition number greater than 30. Condition numbers indicate how sensitive a linear equation solution is to small changes in matrix coefficients. Individual feature results when the condition number is greater than 30 are not included in the variance of the parameter estimates; this impacts standard error diagnostics, global sigma, and standardized residuals. <http://webhelp.esri.com/arcgisdesktop/9.3/>

### 1) 합계출산율의 전역적 공간자기상관 분석(Moran' s I)

본 연구대상지역 전체에 걸쳐 지역별 합계출산율이 공간자기상관성을 가지고 있는지 확인하기 위해 Moran' s I 지수를 산출할 것이다. Moran' s I 지수는 공간자기상관을 파악하기 위한 유용한 척도로서, 지역 간 인접성을 나타내는 공간 가중행렬을 이용하여 이웃하는 지역들이 가지고 있는 속성이 서로 유사한지를 나타내는 통계량이다. 이는 전역적 공간자기상관 분석으로 연구대상 지역에서 유사한 값들의 전반적인 군집 경향을 하나의 지표로 요약하여 나타낸다. 만일 인접한 공간단위들이 연구지역 전체에 걸쳐 서로 유사한 값을 갖는 경우, Moran' s I 값은 양(+)의 값을 갖게 되고 이는 지역 간의 정적인(+) 공간상관을 갖는다는 것을 의미한다. 반대로 연구지역 전반에 걸쳐 인접한 공간 간의 값이 서로 상이한 값을 갖는 패턴을 가진다면 Moran' s I 값은 음(-)의 값을 가질 것이고, 이는 부적인(-) 공간상관을 갖는다는 것을 의미한다. 또한 Moran' s I 값이 0일 경우 공간에 관계없이 무작위 패턴을 가지고 있음을 의미한다. 아래 그림 3을 보면 Moran' s I 값에 따른 공간자기상관과 공간패턴에 대해 쉽게 파악할 수 있다.

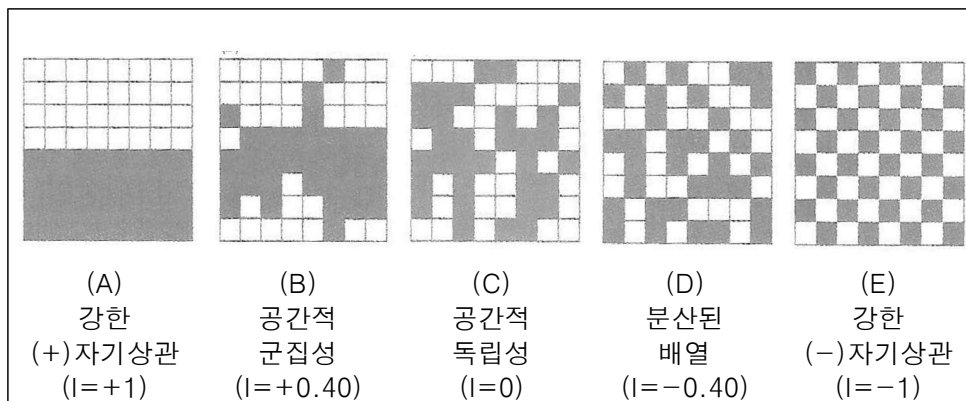


그림 3. Moran' s I 값에 따른 공간자기상관성



Moran' s I 지수는 아래 식 (1)에 의해 계산될 수 있다. 본 연구의 지역별 출산율 분석을 고려한다면 N은 전체 시군구 252개 지역의 개수이고,  $Y_i$  는 i 지역의 출산율이고,  $Y_j$ 는 j지역의 출산율이다.  $\bar{Y}$ 는 전체 252개의 연구대상지역 출산율 평균이고,  $w_{ij}$ 는 공간 인접 가중행렬이다.

$$I = \frac{N \sum_{i=1}^n \sum_{j=1}^n w_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{(\sum_{i=1}^n \sum_{j=1}^n w_{ij}) \sum_{i=1}^n (Y_i - \bar{Y})^2} \quad \text{-----} (1)$$

아울러 Moran' s I 지수는 Z-Score를 통하여 그 통계적 유의성을 검증할 수 있다. Z-score는 아래 식 (2)에 의해 산출될 수 있으며, 이를 통해 Moran' s I 결과에 대한 통계적 검정이 가능해진다.

$$Z_c = \frac{I - E_I}{\sigma_I} \quad \text{-----} (2)$$

여기서  $I$  = Moran' s I의 통계량이고,  $E_I = -1/(n-1)$ ,  $\sigma_I^2 = E[I^2] - E[I]^2$  이다. 본 연구에서는 전역적 Moran' s I 지수를 활용하여 종속변수인 지역별 합계출산율이 연구지역 전반에 걸쳐 공간자기상관성을 띄고 있는지 확인할 것이다. 이 분석을 위해 ArcGIS 10.4.1을 사용할 것이다.

## 2) 합계출산율의 국지적 공간자기상관 분석(Getis-Ord Gi\*)

앞서 살펴본 Moran' s I 지수는 연구대상 지역별로 출산율이 공간적 독립성을 보이는지 혹은 상관성을 보이는지를 알려주지만, 어떤 특정 지역끼리 출산율에 대한 공간 상관성이 있는지는 알려주지 못한다. 즉, 이

지표로는 높은 출산율이 어떤 지역들을 중심으로 나타나고, 낮은 출산율이 어떤 지역들 간에 일어나는지를 파악할 수 없다. 따라서 지역 단위의 국지적 공간자기상관을 분석하여 높은 출산율 군집을 보이는 지역과 낮은 출산율 군집을 보이는지에 대한 확인이 필요하다. 이를 위해 사용할 통계방법은 Getis and Ord(1995)가 제시한 국지적 공간상관성 지수로서, Getis-Ord  $G_i^*$  값이 큰 값들로 군집을 이루어지는 핫스팟(Hot-Spot)과 작은 값으로 군집이 일어나는 콜드스팟(Cold-Spot)을 확인할 수 있게 된다. 다시 말해 높은 출산율이 군집된 지역(Hot-Spot)과 낮은 출산율의 군집이 일어나는 지역(Cold-Spot)을 확인하여 국지적 공간자기상관을 검정할 수 있게 된다. 이러한 통계량을 Getis-Ord  $G_i^*$ 로 부르며 이는 아래의 식(3)-식(5)와 같은 방법으로 산출된다.

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - \left( \sum_{j=1}^n w_{i,j} \right)^2}{n-1}}} \quad \text{-----} \quad (3)$$

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad \text{-----} \quad (4)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad \text{-----} \quad (5)$$

Getis-Ord  $G_i^*$ 의 통계량은 결과 자체만으로 Z-score을 나타내기 때문에 통계적 유의성 검정을 따로 실시할 필요가 없다. 이 값은 지역별로

출산율이 높은 군집(Hot-Spot)과 낮은 군집(Cold-Spot)이 존재하는지를 분석하여 통계적 유의수준을 보여주며 이를 GIS를 통해 지도에 구현할 수 있도록 한다. 이는 관찰하고자 하는 지역의 관찰값을 이웃해있는 지역의 값과 함께 고려하여 분석하는 방식이다. 상대적으로 상이한 값을 가지는 것만으로 통계적 유의성을 갖는 것이 아니라 관찰하고자 하는 지역의 값이 이웃지역에서도 유사하게 일어나야 통계적 유의성을 확인할 수 있다. 따라서  $G_i^*$ 이 통계적으로 유의한 양(+)의 값을 보이는 곳은 상대적으로 높은 출산율의 군집을 이루는 지역을 나타내며, 통계적으로 유의한 음(-)의 값은 상대적으로 낮은 출산율의 군집을 이루는 지역을 나타낸다.

본 연구는 Getis-Ord  $G_i^*$  분석을 통해 높은 출산율과 낮은 출산율의 군집성을 이루는 지역을 확인하여 출산율의 핫스팟과 콜드스팟 분석을 실시하고, 출산율의 공간자기상관이 국지적으로도 이뤄지고 있음을 분석할 것이다. Getis-Ord  $G_i^*$  분석을 위해 ArcGIS 10.4.1을 활용할 것이고 이를 지도에 구현할 것이다.

#### 4. 지리가중회귀모형 분석

지리가중회귀모형은 지역적 차원에서 회귀계수를 추정하는 것이라고 할 수 있다. 즉 변수들 간의 관계를 추정하는 회귀계수가 지역별로 다르다는 것을 전제하여 지역별로 국지적 회귀모형(Locally Regressed Model)을 추정하는 것이다. 지역별로 회귀계수가 다르다는 점에서 다수준 모델(Multi-Level Model)과 유사하나 지역별 관찰치가 다수준 모델에서는 이산형(discrete)으로 분포되어 있음을 가정하고 지리가중회귀모형은 연속형(continuous)으로 분포되어있음을 가정한다는 점에서 차이가 있다.

회귀계수가 지역별로 다르게 추정되는 것, 즉 공간적 이질성을 분석하기 위해서는 전역적 모델(Global Model)과 국지적 모델(Local Model)에

대한 이해가 필요하다. 전역적 모델은 관측치가 위치해 있는 공간에 따른 효과가 없다는 가정을 두고 분석을 실시하는 것이고, 국지적 모델은 전역적 모델을 세분화하여 공간과 위치에 따른 차이와 효과를 분석하는 모델이다. 전역적 모델의 경우 하나의 고정된 회귀계수가 적용되지만, 국지적 모델은 지역별로 다른 회귀계수가 도출된다(이희연, 심재현, 2013). 표 6은 전역적 모델 중 가장 널리 사용되고 있는 최승자승법(OLS: Ordinary Least Squares)과 지리가중회귀모델을 비교한 것이다.

아래 표 6에 언급된 것과 같이 기존의 OLS 회귀모델은 공간적 동질성과 유사성을 가정하여 통계값을 산출하는 반면, GWR모델은 공간적 이질성을 가정하고 분석을 실시한다. 또한 GWR 모델은 산출된 통계량을 지도화할 수 있기 때문에 시각적으로도 분석결과를 확인할 수 있다는 장점이 있다.

표 6. 전역적 모델, 국지적 모델 비교

| OLS 회귀모델           | 지리가중회귀모델(GWR)        |
|--------------------|----------------------|
| 전체 지역에 대한 분석       | 전역적 통계치에 대한 국지적인 상세화 |
| 단일한 하나의 통계값 산출     | 다양한 통계값 산출           |
| 지도화할 수 없음          | 지도화하기 용이함            |
| 비공간적 또는 공간적 접근의 한계 | 공간적 접근에 기반을 둠        |
| 공간적 동질성 또는 유사성을 강조 | 공간적 이질성을 강조          |

※ 출처: Fotheringham, A. et al. (2002).

위의 차이점을 함수로 설명하면 다음과 같다. 우선 OLS 회귀모형은 기존 연구에서 가장 일반적으로 적용되는 통계모형이다. 이론적으로는 관측치의 패턴이 공간적으로 고정된(stationary) 형태라고 가정한 선형 함수

수형태이며, 식 (7)과 같이 정리할 수 있다.

$$y = \beta_0 + \beta x_1 + \beta x_2 + \dots + \beta x_n \text{ ----- (7)}$$

OLS 회귀모형에서는 회귀계수  $\hat{\beta}$ 을 구하기 위해서는 최소자승법에 따라 식 (8)과 같은 함수를 활용할 수 있다.

$$\hat{\beta} = [X'X]^{-1}X'Y \text{ ----- (8)}$$

그러나 본 연구에서 수행할 공간자기상관성 검정이 통계적으로 유의미한 결과를 갖게 되면 OLS 모형의 적용이 타당하지 않을 수 있음을 의미한다. 따라서 공간 이질성을 고려한 모델인 GWR 모형을 적용할 수 있는데, 이는 식 (7)의 선형 함수에 거리함수를 포함시켜 공간이질성을 반영할 수 있는 회귀분석방법이다. GWR 모형은 회귀계수 산출시 각 관측치가 위치한 공간에 거리 가중치를 적용하여 각기 다른 추정 지점에서 공간의 특성에 따라 다른 회귀계수가 산출되며, 이를 산출하기 위한 방식은 식 (9)와 같다.

$$y_i = \beta_{0i}(u) + \beta_{1i}(u)x_{1i} + \beta_{2i}(u)x_{2i} + \dots + \beta_{mi}(u)x_{mi} + \epsilon_i(u) \text{ ----- (9)}$$

$y_i$  = 종속변수( $i = 1, 2, \dots, n$ ),  $n$ 은 관측치

$x_{mi}$  = m번째 독립변수

$\beta_{mi}$  = m번째 회귀계수(파라미터)

$\epsilon_i$  = 오차항

$u$  = 공간좌표

GWR 모형 역시 회귀계수  $\hat{\beta}$ 을 구하기 위해서는 OLS 방식과 비슷하지만 각  $u$ 의 위치에 대한 거리가중함수를 적용하는 가중최소자승법

(Weighted Least Squares)의 구조를 가진다.

$$\hat{\beta}(u) = [X' W(u) X]^{-1} X' W(u) Y \quad \text{-----} \quad (10)$$

여기서 관측 위치인  $u$  지점에 대한 공간자기상관을 고려한 공간행렬 ( $W$ )과 결합하게 되는데 가까운 지역들의 영향이 먼 지역들의 영향보다 더 크게 고려된다, 식 (11) 과 같이  $W_i(u)$ 는  $i$ 번째 관측값과 다른 모든 관측값들 간의 거리를 반영하는 공간행렬로,  $i$ 번째 관측값의 거리에 기초한 가중치  $d_i$  벡터를 포함하는  $n \times n$  대각행렬이다.

$$W_i(u) = \begin{pmatrix} w_1(u) & 0 & 0 & \dots & 0 \\ 0 & w_2(u) & 0 & \dots & 0 \\ 0 & 0 & w_3(u) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & w_n(u) \end{pmatrix} \quad \text{-----} \quad (11)$$

공간행렬을 구축하는 경우, 두 지역 간의 지리적 가중함수를 어떻게 정의하고 측정하는가는 매우 중요하다. 연구에서 일반적으로 사용되는 지수식 가중치인 가우시안(Gaussian) 방법을 사용할 것이며 이는 식 (12)로 산출할 수 있다.

$$W_i(u) = e^{-0.5(d_i(u)/\theta)^2} \quad \text{-----} \quad (12)$$

여기서  $d$ 는  $i$ 지역에서 다른 지역까지의 거리이며,  $\theta$ 는 대역폭(Bandwidth)이라고 불리는 값이다. 따라서 그림 4에서 볼 수 있는 바와

같이 대역폭의 설정에 따라 가중치는 달라진다.  $\theta$ 가 커질수록 동일한 거리에 대한 가중치는 커지게 되며,  $\theta$ 가 작아질수록 동일한 거리에 대한 가중치 값은 0에 근접하게 된다. 대역폭이 증가하여 공간 가중치가 1에 가깝게 되면 지리가중회귀모델은 OLS 회귀모델과 유사하게 된다(이희연, 심재현, 2013).

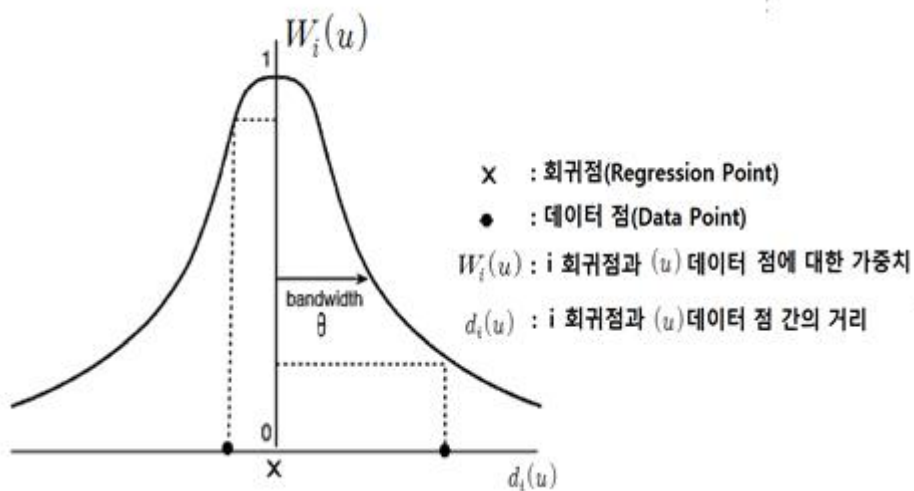
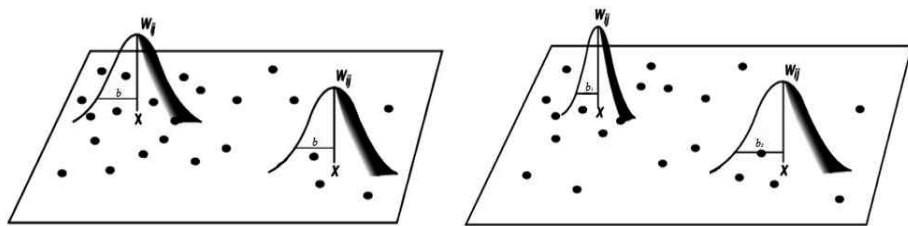


그림 4. 지리가중회귀모형의 공간가중 개념

※ 출처: Fotheringham, A. et al.(2002)

앞서 언급한대로 지리가중회귀모델은 대역폭( $\theta$ : Bandwidth)의 영향을 많이 받기 때문에 적절한 대역폭을 설정하여 모델의 적합도를 높여야 한다. 특히 대역폭을 고정된(fixed) 방식으로 설정할 것인가 혹은 조정된(adaptive)방식으로 설정할 것인가를 결정해야 하는데, 연구대상지역에서 관측치들의 위치가 규칙적으로 분포하는 경우에는 고정된 대역폭 추정방식을 사용할 수 있다. 하지만 관측치들의 위치가 불규칙하게 흩어져 있는 경우에는 고정된 대역폭으로 추정하게 되면 모델의 적합도가 떨어진다.

불규칙적인 관측치들의 분포가 예상된다면 조정된 대역폭 추정방식을 선정하여 모델의 적합도를 증가시킬 수 있다. 조정된 대역폭 설정방법은 관측치들의 위치가 몰려있는 지역에서는 대역폭을 작게 하고, 그렇지 않은 지역에서는 대역폭을 넓게 하는 방식이다. 고정된 대역폭 추정방식과 조정된 대역폭 추정 방식의 비교는 그림 5를 통해 이해할 수 있다.



(a) 고정된(fixed) 대역폭 설정

(b) 조정된(adaptive) 대역폭 설정

그림 5. 대역폭(Bandwidth) 설정 방법 비교

※ 출처: Fotheringham, A. et al.(2002)

적정한 대역폭의 설정에 관한 검정은 관찰값과 추정값 간의 차이인 교차검증(CV: Cross Validation)값을 최소화하는 방법으로 선정할 수 있으며, CV 산출식은 다음 식 (13)과 같다.

$$CV = \sum [Y_i - \hat{Y}_i(\theta)]^2 \text{ ----- (13)}$$

그러나 관찰값과 추정값의 차이만이 아니라 모델의 적합성을 판정하는데도 사용되는 AIC(Akaike Information Criterion)가 가장 많이 이용된다. 대역폭의 설정은 자동적으로 계산되게 하거나 직접 입력하는 방법이 있는데, 사전적으로 대역폭을 지적할만한 충분한 근거가 없는 경우에는 자동적으로 계산되도록 하는 것이 관행이다(Fotheringham, et al., 2008).



AIC는 설정되는 회귀점으로부터 어느 정도 대역폭에 있는 표본점들을 선택할 것인가를 판정하고, OLS 모델에 비해 지리가중회귀모델이 보다 더 적합한가를 판정해주는 지표가 된다. AIC는 다음과 같은 식에 따라 계산된다.

$$AIC = 2n\text{Ln}(\hat{\sigma}) + n\text{Ln}(2\pi) + n[n + \text{Tr}(s)]/[n - 2 - \text{Tr}(s)] \quad \text{---} \quad (14)$$

여기서  $n$ 은 데이터의 관측치 개수이며,  $\hat{\sigma}$ 은 잔차들(Residuals)의 표준편차이고,  $\text{Tr}(s)$ 는 지리가중회귀모델의 추정된 매트릭스를 추적한 것이다(Charlton, M., 2009). AIC는 동일한 종속변수에 대해 다른 독립변수들로 구성된 모델을 비교하는데 활용될 수 있으며, 전역적 모델과 국지적 모델을 비교하는 데에도 사용된다.

본 연구에서는 지리가중회귀모형을 적용하고 대역폭 설정을 위해 고정된 방법의 경우 미묘한 공간적 차이를 포착하지 못하여 추정치의 분산을 증가시킬 위험이 있으므로(이희연, 심재현, 2013) 조정된(adaptive) 방법을 사용할 것이다. 또한 대역폭의 값을 검증하기 위한 방법으로 AIC를 하고, OLS모형과 비교를 통해 모형의 적합도를 함께 판정할 것이다. 지리가중회귀모델의 최종 연구모형은 상관분석을 통해 다중공선성이 일어나지 않는 변수를 선정할 것이다. 또한 지리가중회귀모형을 통해 도출된 회귀계수와 결정계수, 표준잔차의 결과를 지도화하여 시각적으로 확인할 수 있도록 할 것이다. 지리가중회귀모형 분석에는 ArcGIS 10.4.1.을 사용할 것이다.

표 7. 연구목적 및 연구방법

| 구분              | 연구목적                            | 연구방법  | 연구대상                             |
|-----------------|---------------------------------|---|----------------------------------|
| 연구<br>문제<br>I   | 지역별 출산율<br>공간패턴 분석              | - 2010년, 2015년<br>시군구 출산율 시각화                     | 2010년,<br>2015년<br>전국<br>252개시군구 |
| 연구<br>문제<br>II  | 지역별 출산율 및 관련<br>변수 공간자기상관<br>검정 | - Moran' s I 지수 도출<br>- Getis-Ord Gi* 분석          |                                  |
| 연구<br>문제<br>III | 지역별 차별출산율에<br>영향을 미치는<br>요인 분석  | - 지리가중회귀모형 이용<br>- OLS 모형과 결과 비교<br>- 지역별 회귀계수 도출 |                                  |

## 제4장 연구결과

### 제1절 연구대상지역의 일반적 특성 및 기본통계 분석결과

문헌고찰을 통해서 지역별 차별출산력과 관련이 있다고 고려된 변수 중 통계청 국가통계포털사이트에서 시군구 별 데이터 수집이 가능한 변수를 선정하여 기초통계량 분석 후 지역별 현황을 파악하였다. 본 연구의 종속변수인 출산율은 2010년 1.23, 2015년 평균 1.24로 비슷한 수준을 보였다. 2010년 출산율이 가장 높은 곳은 전라북도 진안군 (2.41)이었고, 2015년 출산율이 가장 높은 곳은 전라남도 해남군(2.46)이었다. 출산율이 가장 낮은 곳은 2015년은 서울시 종로구(0.813), 2010년 부산 서구(0.78)이었으며 대체적으로 서울과 부산지역에서 낮은 출산율이 기록되었다.

조혼인율은 2010년 1000명당 평균 5.9명이었으며 2015년에는 5.4명으로 다소 줄었으나 큰 차이는 없었다. 인구밀도의 경우 지역별 편차가 매우 컸으며, 2010년의 경우 인구밀도가 가장 낮은 곳은 강원도 인제군이었으며 양천구의 인구밀도가 가장 높았다. 2015년에는 가평군의 인구밀도가 가장 낮았으며 부산 남구의 인구밀도가 가장 높았다. 하지만 2015년의 인구밀도 편차는 줄어들었다. 여성의 초혼연령은 2010년 평균 28.89였으며 2015년에는 29.85명으로 약 1세가량 올랐다. 가임기 여성비율은 전체인구 중 가임기(15세-49세)여성의 비율을 나타내는 것으로 실제 출산이 가능한 인구비율을 말한다. 이는 2010년 평균 약 0.54이었으나, 2015년 0.43으로 줄어들었다. 가임기 여성비율이 낮은 곳은 대체로 고령인구비율이 높은 전라남도에 위치해 있었다. 외국인과의 혼인율은 전체 혼인건수 중 한국인 남편과 외국인 부인이 혼인한 건수의 비율을

나타내는 것으로 2010년 약 9% 이었으나 2015년에는 6%로 줄어들었다. 외국인 혼인율이 가장 높은 곳은 2015년 경상남도 거창군이었다. 여성의 취업인구비율은 2010년 평균 0.48이었으며 2015년 0.46으로 비슷한 수준이었다. 아파트 전세값(만원/ $m^2$ )의 경우 지역별 편차가 매우 컸으며 2010년 최대값이 1 제곱미터당 437만원이었으나 2015년에는 607만원으로 올랐으며 지역편차도 70에서 106으로 크게 올랐다. 2015년 전세값의 평균은 166만원이었으며 중앙값은 139만원이었으나 전세값이 가장 낮은 곳(전라북도 순창군: 29만원)과 가장 높은 곳(서울시 강남구: 607만원)의 차이가 컸다. 인구 1000명당 사업체 수 역시 지역별 편차가 컸으며 사업체 수가 가장 많은 곳은 2010년 2015년 모두 서울시 중구였다. 유아(0-4세) 1000명당 보육시설 수는 2010년 평균 약 15개였으며 2015년 약 17개로 증가했다. 또한 보육시설이 제일 적었던 지역은 유아 천명당 약 2곳이었는데, 2015년에는 보육시설이 제일 적은 곳의 보육시설 수는 약 7곳으로 증가했다.

고령인구비율이 가장 높은 곳은 2010년 2015년 모두 고흥군이었고, 대부분 전라도 지역이 고령인구 비율이 높은 것으로 나타났다. 노인(60세 이상) 천 명당 노인복지시설은 2010년과 2015년 평균 약 10곳이었다. 표 8과 표 9는 연구변수의 기초통계량을 정리한 것이고 그림 6은 전국 합계출산율을 GIS를 통해 시각화한 것이다.

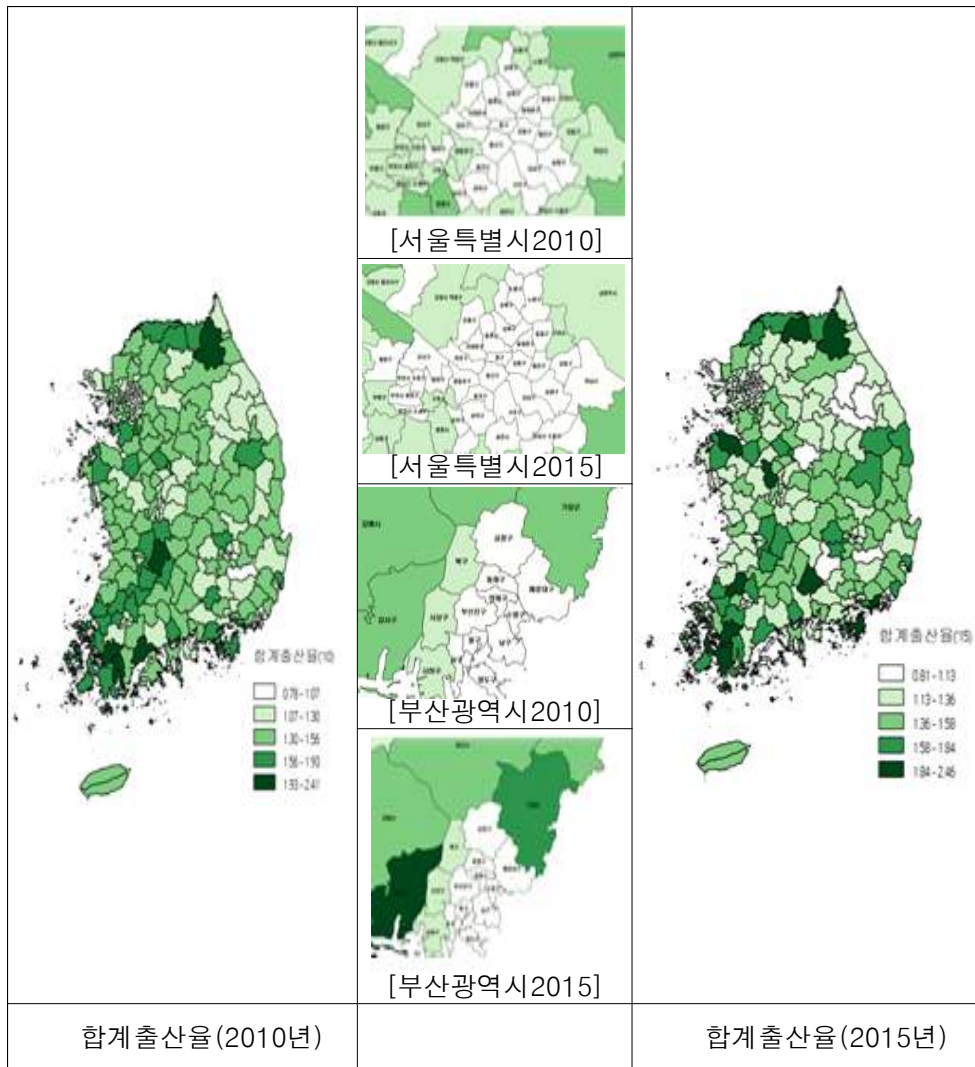


그림 6. 2010년, 2015년 지역별 합계출산율

표 8. 연구변수 기초통계량(2010년도)

| 지표명   |                           |                                | 최소값   | 최대값    | 중앙값   | 평균     | 표준편차  |
|-------|---------------------------|--------------------------------|-------|--------|-------|--------|-------|
| 종속 변수 | 합계출산율                     |                                | 0.78  | 2.41   | 1.31  | 1.23   | 0.26  |
|       |                           |                                |       |        |       |        |       |
| 독립 변수 | 인구<br>·<br>결혼             | 조혼인율(명)                        | 3.80  | 10.30  | 6.00  | 5.96   | 1.19  |
|       |                           | 인구밀도(명)                        | 20    | 29028  | 588   | 4334   | 6551  |
|       |                           | 인구증가율(%)                       | -3.49 | 14.81  | 0.85  | 1.25   | 2.14  |
|       |                           | 여성<br>초혼연령(나이)                 | 27.68 | 30.22  | 28.90 | 28.89  | 0.61  |
|       |                           | 가임기여성비율                        | 0.20  | 0.80   | 0.58  | 0.54   | 0.14  |
|       |                           | 가임기여성<br>미혼율                   | 0.06  | 0.38   | 0.23  | 0.21   | 0.07  |
|       |                           | 외국인과의<br>혼인율                   | 0.03  | 0.27   | 0.09  | 0.1    | 0.05  |
|       | 여성                        | 여성<br>취업인구비율                   | 0.37  | 0.68   | 0.47  | 0.48   | 0.06  |
|       |                           | 대졸이상<br>여성 인구비율                | 0.03  | 0.56   | 0.24  | 0.25   | 0.09  |
|       | 지역<br>경제<br>·<br>출산<br>환경 | 아파트<br>전세가격(만원)                | 21.00 | 437.32 | 98.55 | 110.98 | 70.30 |
|       |                           | 인구 1000명당<br>사업체수              | 40.07 | 436.04 | 66.39 | 71.23  | 34.03 |
|       |                           | 유아 1,000명당<br>보육시설 수           | 1.40  | 30.00  | 14.80 | 15.37  | 4.58  |
|       | 노인                        | 고령인구비율                         | 0.05  | 0.31   | 0.12  | 0.15   | 0.73  |
|       |                           | 노인(60세이상)<br>1000명당<br>노인복지시설수 | 1.70  | 34.70  | 8.80  | 10.80  | 7.80  |
|       |                           | 건강보험료<br>대비 급여비용               | 0.06  | 11.29  | 2.17  | 2.41   | 1.39  |

표 9. 연구변수 기초통계량(2015년도)

| 지표명  |                           |                                 | 최소값   | 최대값    | 중앙값   | 평균    | 표준편차   |
|------|---------------------------|---------------------------------|-------|--------|-------|-------|--------|
| 종속변수 | 합계 출산율                    |                                 | 0.81  | 2.46   | 1.20  | 1.24  | 0.26   |
|      |                           |                                 |       |        |       |       |        |
| 독립변수 | 인구<br>·<br>결혼             | 조혼인율(명)                         | 2.97  | 9.31   | 5.32  | 5.40  | 1.16   |
|      |                           | 인구밀도(명)                         | 17    | 26095  | 575   | 3924  | 5932   |
|      |                           | 인구증가율(%)                        | -3.05 | 15.51  | 0.06  | 0.39  | 1.95   |
|      |                           | 여성<br>초혼연령(나이)                  | 28.08 | 31.46  | 29.82 | 29.85 | 0.65   |
|      |                           | 가임기여성비율                         | 0.22  | 0.61   | 0.47  | 0.43  | 0.09   |
|      |                           | 가임기여성<br>미혼율                    | 0.11  | 0.38   | 0.24  | 0.23  | 0.06   |
|      |                           | 외국인과의<br>혼인율                    | 0.02  | 0.20   | 0.05  | 0.06  | 0.03   |
|      | 여성                        | 여성<br>취업인구비율                    | 0.36  | 0.63   | 0.46  | 0.46  | 0.04   |
|      |                           | 대졸이상<br>여성 인구비율                 | 0.19  | 0.69   | 0.32  | 0.33  | 0.08   |
|      | 지역<br>경제<br>·<br>출산<br>환경 | 아파트<br>전세가격(만원)                 | 29    | 607    | 139   | 166   | 106.19 |
|      |                           | 인구 1000명당<br>사업체수               | 44.77 | 510.40 | 75.39 | 80.58 | 39.41  |
|      |                           | 유아 1,000명당<br>보육시설 수            | 7.38  | 31.79  | 16.92 | 17.25 | 4.34   |
|      | 노인                        | 고령인구비율                          | 0.06  | 0.36   | 0.15  | 0.18  | 0.76   |
|      |                           | 노인(60세 이상)<br>1000명당<br>노인복지시설수 | 1.36  | 32.70  | 8.03  | 10.09 | 7.72   |
|      |                           | 건강보험료<br>대비 급여비용                | 0.40  | 4.56   | 1.45  | 1.78  | 0.95   |

## 제2절 상관관계 분석결과

종속변수인 합계출산율과 독립변수들에 대한 단순상관분석 결과, 표 10과 같이 조혼인율을 제외하고 모든 인구·결혼, 여성, 지역경제와 출산환경, 노인 관련 변수에서 통계적으로 유의미한 상관관계가 있는 것으로 나타났다.

표 10. 독립변수와 출산율과의 상관관계(2010, 2015년도)

| 변수명               |                      | 상관계수<br>(2010) | 상관계수<br>(2015) |
|-------------------|----------------------|----------------|----------------|
| 인구<br>·<br>결혼     | 조혼인율                 | 0.006          | 0.012          |
|                   | 인구밀도(log)            | -0.17**        | -0.16**        |
|                   | 인구증가율                | 0.016*         | 0.044*         |
|                   | 여성 초혼연령              | -0.26**        | -0.19**        |
|                   | 가임기여성비율              | -0.59**        | -0.83**        |
|                   | 가임기여성 미혼율            | -2.001**       | -2.273**       |
|                   | 외국인과의 혼인율            | 0.018**        | 0.023**        |
| 여성                | 여성 취업인구비율            | 1.52**         | 1.44**         |
|                   | 대졸이상 여성 인구비율         | -1.36**        | -1.28**        |
| 지역경제<br>·<br>출산환경 | 아파트 전세가격(log)        | -0.516**       | -0.584**       |
|                   | 인구 1000명당 사업체 수(log) | -0.001**       | -0.001**       |
|                   | 유아 1000명당 보육시설 수     | -0.008*        | -0.013**       |
| 노인                | 고령인구비율               | 0.013**        | 0.008**        |
|                   | 노인 1000명당 노인복지시설 수   | 0.21**         | 0.017**        |
|                   | 건강보험료대비 급여비용         | 0.007**        | 0.088**        |

※ 주 : \* p<0.05, \*\* p<0.01



출산율과 양(+의 상관관계가 있는 변수는 인구증가율, 여성취업인구 비율이었으며, 특히 노인과 관련된 모든 변수에서 통계적 유의미한 양(+의 관계가 있는 것으로 나타났다. 단순 상관관계만 보았을 때, 노인인구의 증가가 출산율에 단순히 부정적인 영향만을 주는 것이 아님을 확인할 수 있다. 예를 들면 노인인구가 상대적으로 많은 지역에서 출산에 대한 전통적인 가치가 존재하거나, 양육지원효과가 있을 수 있다. 이를 제외한 다른 모든 변수에서는 출산율과 음(-)의 영향을 보였다. 출산율과 변수간의 양/음의 관계는 2010년과 2015년에 동일한 것으로 나타났고, 이는 5년 동안 출산율과 변수간의 상관관계에 큰 변화가 없었던 것을 의미한다.

### 제3절 공간자기상관 분석결과

제2장 연구방법에서 언급한 것과 같이 공간자기상관 검정을 위해 두 가지 분석을 실시하였다. 첫 번째 분석은 전국적인 공간자기상관성을 분석하기 위하여 Moran's I 지수를 활용하고 전국 시군구별 합계출산율과 관련변수에 대한 공간적 자기상관성을 검정하였다. 두 번째 분석은 지역별 공간자기상관성을 분석하기 위해 Getis-ord  $G_i^*$  지수를 활용하여 특정 지역 간의 공간상관성을 통해 출산율과 관련변수가 특정지역에서 군집이 형성되는지 검정할 것이다. 이 두 가지 분석에 대한 결과는 아래와 같다.

#### 1. 지역별 출산율과 관련변수의 전역적 공간자기상관 검정

시군구별 합계출산율의 공간자기상관을 검정하기 위해 Moran's I 지수를 통한 분석을 실시한 결과, 표 11과 같은 결과를 도출하였다. 합계출

산율의 경우 2010년 Moran 지수가 0.32에서 2015년 0.34로 다소 증가하였으며, 이는 공간적인 자기상관성이 다소 강해졌음을 의미한다. 조혼인율의 경우 Moran 지수가 2010년 0.31에서 2015년 0.26으로 떨어져 혼인율에 대한 공간적 자기상관이 약화되었음을 알 수 있다. 인구밀도의 경우 Moran' I 가 2010년 0.63에서 2015년 0.73으로 증가하여 인구 밀도에 대한 공간자기상관성이 매우 높음을 알 수 있다. 이로 인해 특정지역에 인구가 몰리고 있는 현상을 추측할 수 있다. 여성의 초혼연령은 2015년이 2010년에 비해 Moran 's I 지수가 다소 감소하였으나 2015년 0.4를 보이며 특정지역에 비슷한 수준의 초혼연령대가 몰려있음을 예측해볼 수 있다. 마찬가지로 가임기여성의 미혼율 경우 2010년, 2015년 두해 모두 0.5 이상의 Moran' I가 나타나, 미혼의 가임기여성이 특정지역에 밀집해 있는 것으로 해석할 수 있다. 외국인과 혼인의 경우 0.2수준의 Moran' s I를 나타내며 외국인과의 혼인이 지역적으로 모여 있지 않다는 것을 알 수 있다.

여성관련 변수를 살펴보면 여성취업인구 비율은 2010년 0.27에서 2015년 0.13으로 여성취업인구비율은 공간적 독립성이 존재함을 알 수 있다. 반면, 대졸이상 여성비율은 0.3~0.4로 특정지역에 여성 대학졸업자 비율이 밀집되어 있는 것으로 나타났다. 지역경제변수를 살펴보면 아파트 전세가격은 2010년 0.63에서 2015년 0.89으로 매우 높아졌으며, 이는 전세가격이 특정지역에만 높게 측정되어 있음을 의미한다. 반면 인구 1000명당 사업체 수는 지역 간의 지리적인 영향이 적은 것으로 나타났다.

노인과 관련된 변수를 살펴보면 고령인구 비율과, 노인복지시설수, 보혈료대비 급여비 비율 모두 양(+)의 공간군집성을 나타내는 것으로 나타났다. 이는 특정지역에 고령인구가 모여 있기 때문에 노인복지시설과 급여비 지출이 높은 지역이 유사할 것이라는 것을 예측해 볼 수 있다. 마

지막으로 유아 1000명당 보육시설 수는 공간적인 관계가 매우 적은 것으로 나타났다.

표 11. 시군구 합계출산율에 대한 Moran' s I 분석 결과

| 변수명               |                      | 2010년도<br>Moran's I | 2015년도<br>Moran's I |
|-------------------|----------------------|---------------------|---------------------|
| 합계출산율             |                      | 0.32                | 0.34                |
| 인구<br>·<br>결혼     | 조혼인율                 | 0.31                | 0.26                |
|                   | 인구밀도(log)            | 0.63                | 0.73                |
|                   | 인구증가율                | 0.07                | 0.06                |
|                   | 여성 초혼연령              | 0.58                | 0.40                |
|                   | 가임기여성비율              | 0.37                | 0.40                |
|                   | 가임기여성 미혼율            | 0.53                | 0.56                |
|                   | 외국인과의 혼인율            | 0.27                | 0.21                |
| 여성                | 여성 취업인구비율            | 0.27                | 0.13                |
|                   | 대졸이상 여성 인구비율         | 0.41                | 0.34                |
| 지역경제<br>·<br>출산환경 | 아파트 전세가격(log)        | 0.63                | 0.89                |
|                   | 인구 1000명당 사업체 수(log) | 0.03                | 0.02                |
|                   | 유아 1000명당 보육시설 수     | 0.06                | 0.11                |
| 노인                | 고령인구비율               | 0.36                | 0.34                |
|                   | 노인 1000명당 노인복지시설 수   | 0.48                | 0.52                |
|                   | 건강보험료대비 급여비용         | 0.33                | 0.42                |

앞서 살펴 본 바와 같이 지역별 합계출산율과 관련 변수들을 Moran' s I를 통해 분석해 본 결과, 대부분의 변수에서 공간자기상관

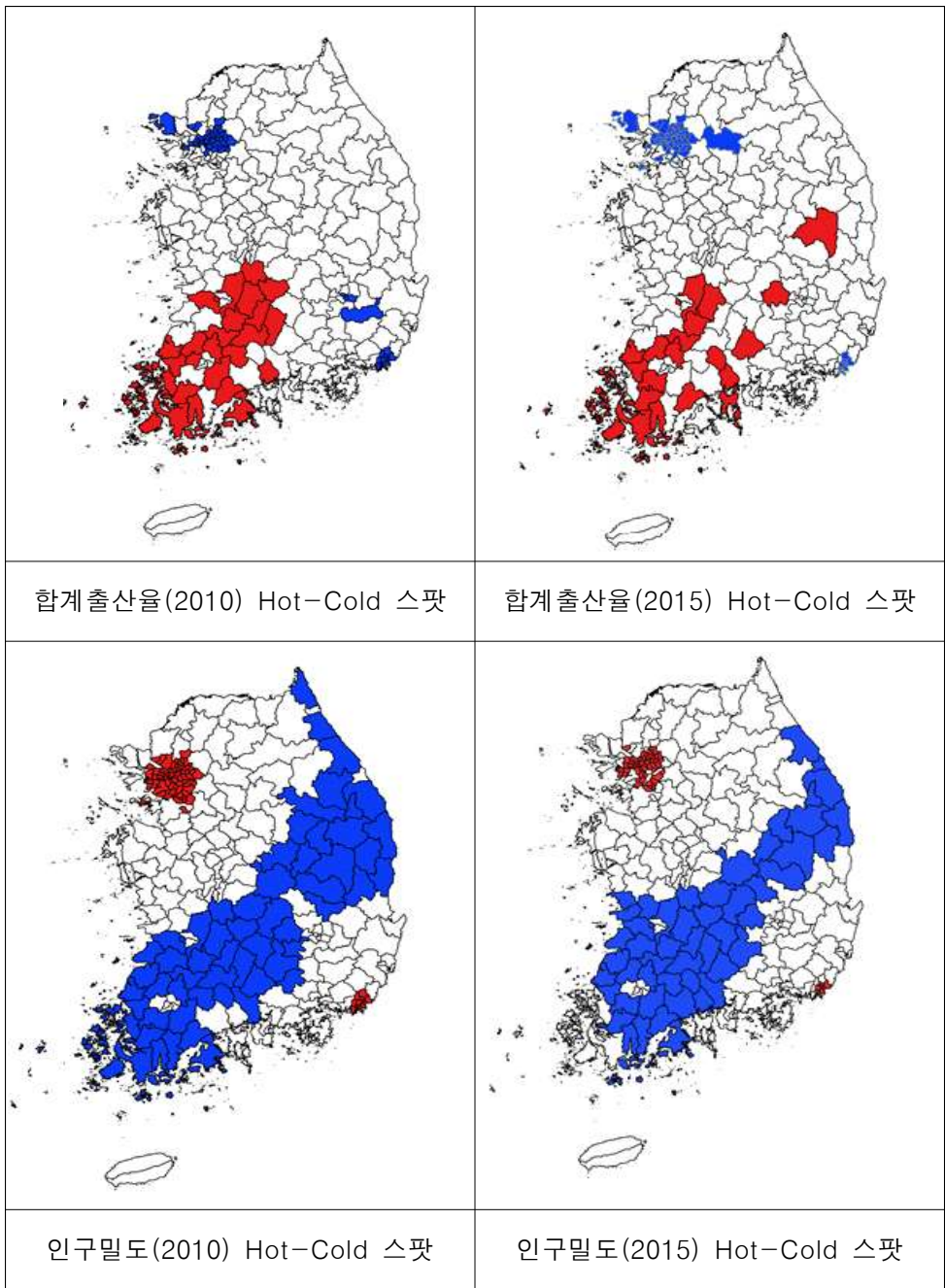
(Spatial Autocorrelation)이 존재하는 것으로 확인되었다. 하지만 Moran 지수는 전국적인 자기상관을 확인할 수 있지만, 어느 지역에서 유사한 수준의 변수들이 밀집해 있는지 확인을 할 수 없다. 이를 확인하기 위하여 국지적 자기상관을 확인할 수 있는 Getis-Ord  $G_i^*$ 을 확인하여 출산율과 관련 변수들이 어느 지역에서 유사한 값으로 밀집해 있는지 확인해보겠다. 만약 높은 값끼리 밀집해 있는 지역이 있다면 그 지역은 해당 변수의 핫스팟이고 낮은 값끼리 모여 있다면 그 지역은 특정 변수의 콜드스팟으로 볼 수 있다. 이로 인해 Getis-Ord  $G_i^*$  분석은 핫스팟 분석(Hot-Spot Analysis)으로 불리기도 한다. 본 연구에서는 Moran' I 분석을 통해서 2010년, 2015년 두 해의 Moran 지수값이 0.3이상인 변수들만 선정하여 Getis-Ord  $G_i^*$  분석을 실시하겠다. 0.3이하인 경우 공간적 군집현상이 일어나기는 하지만 군집현상이 매우 미비하게 일어나기 때문에 본 연구에서는 이를 제외하도록 하겠다.

## 2. 지역별 출산율과 관련변수의 국지적 공간자기상관 검정

### 1) 인구·결혼 변수 Getis-Ord $G_i^*$ 분석결과

합계출산율과 함께 인구·결혼관련 변수 중 2010년과 2015년 두 해에 걸쳐 0.3이상의 Moran' s I 값을 보인 인구밀도, 여성초혼연령, 가임기여성비율, 가임기여성 미혼율에 대한 분석을 실시하도록 하겠다. 그림 7은 인구·결혼 변수에 대한 Getis-Ord  $G_i^*$  분석결과를 지도화 한 것이다. 지도에서 빨간색으로 표시된 곳은 변수가 높은값 끼리 특정지역에 밀집해있는 핫스팟이고 파란색으로 표시된 곳은 변수가 낮은값끼리 모여있는 콜드스팟을 나타낸다. 본 연구에선 분석결과를 95%이상의 통계적 유의성을 보이는 지역으로 제한하였다. 인구·결혼 변수에 대한 분석결과를 살펴보면 전체적으로 핫스팟과 콜드스팟의 지역이 유사하게 나타나지

만 시간에 따라 군집현상이 약화되거나 강화되는 변화를 볼 수 있다. 이는 앞서 살펴본 Moran' I의 값이 시간에 따라 감소되거나 증가하는 결과와 같은 것이다. 합계출산율의 경우 전반적으로 서울지역과 부산지역에서 낮은 출산율이 모여 있는 콜드스팟을 형성하고 있고, 전라남도 지역을 중심으로 핫스팟을 형성하는 것을 확인할 수 있다. 인구밀도의 경우 서울과 부산지역에서 핫스팟이 나타나는 것을 볼 수 있고, 전남지역부터 경상남도 지역까지 띠를 이루며 콜드스팟이 형성되어 있는 것을 볼 수 있다. 특히 2010년에 인구밀도의 콜드스팟이 약간 분산되어 있던 것이 2015년에 와서는 확실한 띠 형식의 콜드스팟을 확인할 수 있다. 이것은 앞서 보았듯 Moran' s I 값이 2010년 0.63에서 2015년 0.73으로 증가한 것을 보여주는 것이다. 여성 초혼연령은 서울과 부산지역을 중심으로 초혼연령이 높은 지역끼리 모여 있는 것을 확인할 수 있고, 2010년 전남과 충남지역에서 콜드스팟이 강하게 나타났으나 2015년에 다소 분산되었다. 가임기 여성비율과 가임기여성 미혼율에 대한 Getis-Ord Gi\* 결과는 인구밀도의 결과와 비슷한 양상을 보이는데, 이는 인구밀도에 따라 지역 인구수 많거나 적기 때문일 것으로 추측된다. 출산율과 인구·결혼 변수에 대한 Getis-Ord Gi\* 분석결과의 특징은 서울과 부산에서는 비슷한 결과가 나타나고, 전라남도에서 출산 관련변수에 대해 지속적으로 콜드스팟이 확인된다는 것이다. 이는 대도시와 지방도시의 특성상 나타나는 현상으로 해석할 수도 있다. 이는 도시지역에서는 지속적으로 낮은 출산율이 확대되고, 지방지역에서는 상대적으로 높은 출산율이 지속적으로 유지된다고 말할 수 있다.



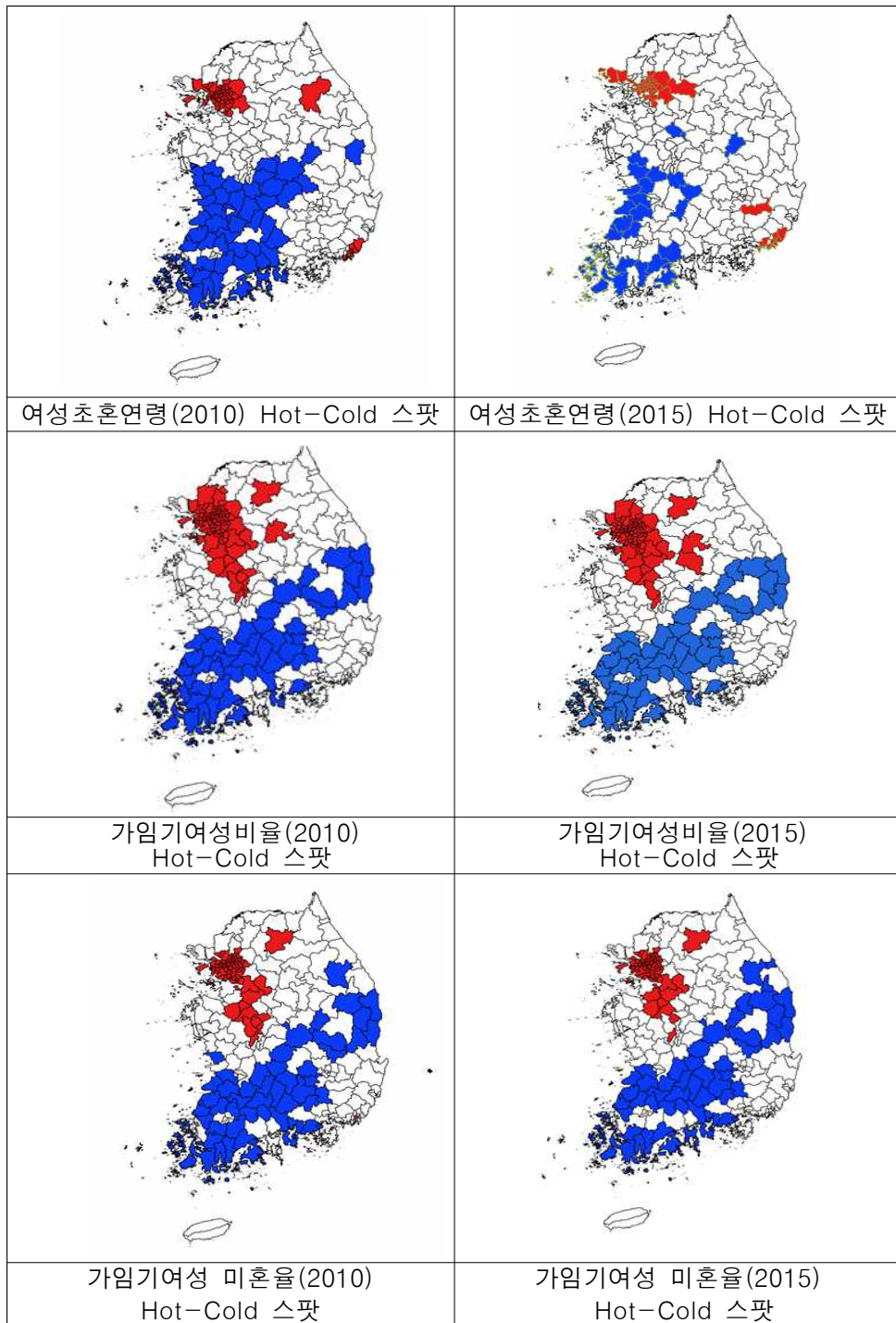


그림 7. 인구변수 Getis-Ord Gi\* 분석결과

## 2) 여성 관련변수 Getis-Ord Gi\* 분석결과

여성관련 변수에 대한 Moran' s I 분석결과 2010년, 2015년 두 해에 걸쳐 값이 0.3 이상 이었던 변수는 대졸이상 여성인구(25세-49세)비율이 었다. 그림 8은 고학력 여성인구 비율에 대한 Getis-Ord Gi\* 분석결과를 나타낸 것이다. 이는 앞서 본 지역별 출산율 Getis-Ord Gi\* 분석결과와 반대의 결과를 보이고 있다. 서울과 부산에서 대졸 이상의 여성비율이 높은 지역이 밀집해있기 때문에 핫스팟 현상을 보이며 전라남도 지역과 충청·대전지역에는 상대적으로 낮은 대졸이상의 여성비율이 밀집되어 콜드스팟을 보인다. 이는 여성 교육수준이 높을수록 기회비용에 따른 출산율이 감소한다는 합리적인 이론을 반영한다고 할 수 있다. 지역적으로 살펴본다면 서울과 부산지역과 같은 대도시에선 대졸이상의 여성인구비율이 높아 출산율이 낮고, 지방도시에서는 대학졸업 이상의 여성인구 비율이 낮아 출산율이 상대적으로 높을 것으로 추측해 볼 수 있다.

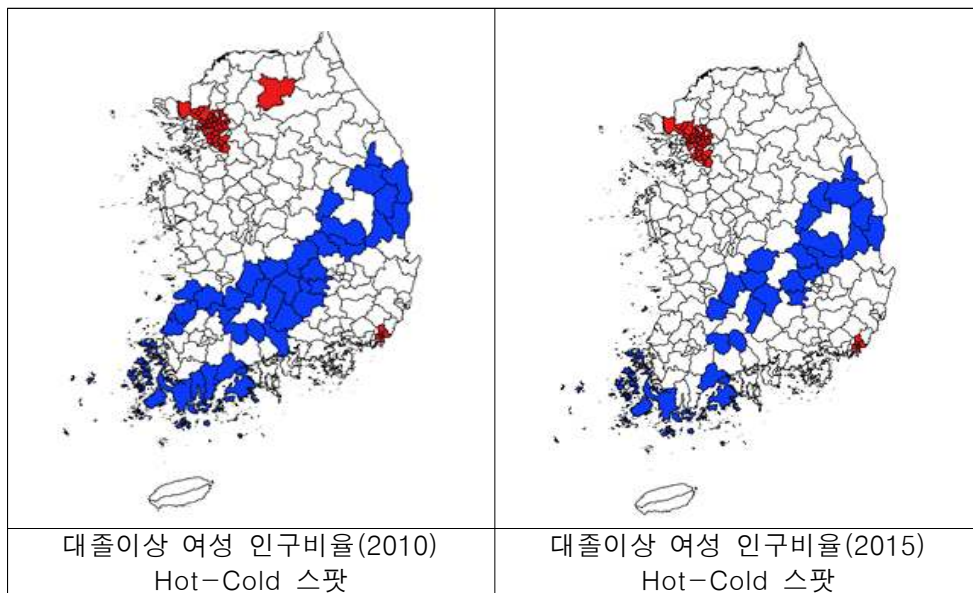


그림 8. 여성 관련변수 Getis-Ord Gi\* 분석결과



### 3) 지역경제 변수 Getis-Ord Gi\* 분석 결과

Moran's I 분석결과 2010년, 2015년 두 해 모두 값이 0.3이상 이었던 지역경제 변수는 아파트 전세값이었다. 그림 9는 아파트 전세값에 대한 Getis-Ord Gi\* 분석결과를 나타낸 것이다. 앞서 살펴본 인구밀도의 결과와 매우 유사한 결과를 보이고 있으며, 이는 인구밀도가 높은 곳의 아파트전세가격이 높다는 것을 파악할 수 있다. 2015년 아파트 전세값은 2010년도 보다 공간적인 상관성이 더욱 강해지는 것을 확인할 수 있다. 출산율의 결과와 반대의 핫스팟 지역과 콜드스팟 지역을 보이기 때문에 인구밀도와 전세값이 높을수록 출산율이 낮아진다는 것을 확인할 수 있다.

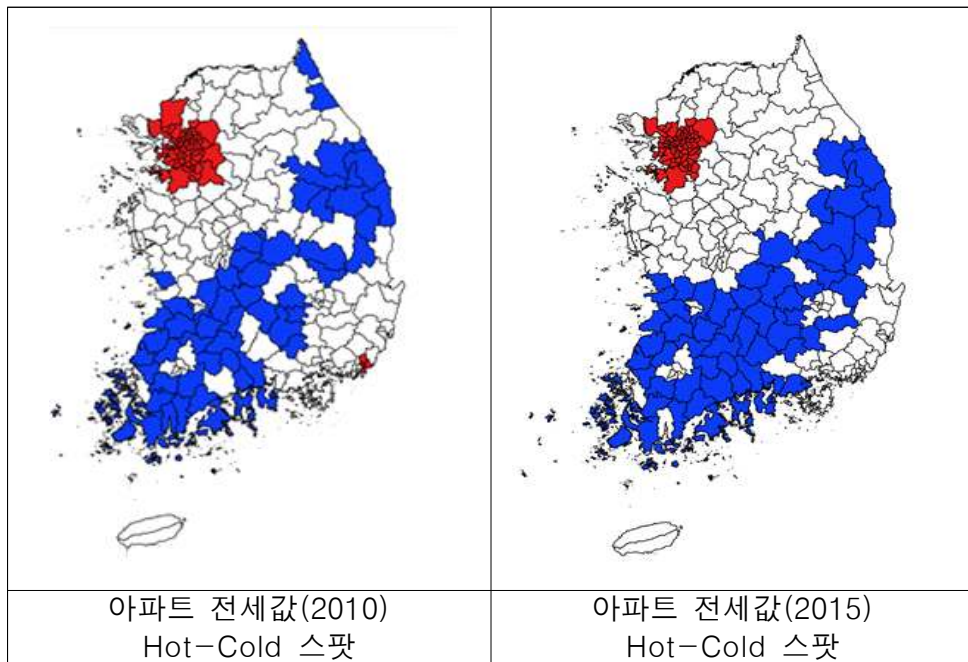


그림 9. 지역경제변수 Getis-Ord Gi\* 분석결과

#### 4) 노인 변수 Getis-Ord $G_i^*$ 분석결과

Moran's I 분석결과 노인변수는 3개 변수 모두 2010년, 2015년 모두 0.3 이상의 Moran 지수값을 보였다. 그림 10은 노인 변수에 대한 분석 결과를 나타낸 것이다. 고령인구 비율은 서울과 수도권에서 낮게 나타났으며 유사한 수준의 지역이 밀집해 있는 것으로 나타났고, 전남, 대전, 경북지역에서 고령인구비율이 높은 콜드스팟 지역이 확인되었다. 노인 1000명당 노인복지시설은 노인인구비율이 높은 지역에 많이 설치되어 운영이 되는 것을 예측할 수 있고, 지역 건강보험금액 대비 급여비용 비율도 고령인구비율의 분석결과와 양상이 매우 유사하였다. 주목할 부분은 앞서 살펴본 인구·결혼, 여성, 지역경제 결과와 반대의 결과가 나오는 것으로 보아 인구밀도가 낮고, 아파트 전세값이 낮은 지역에 노인인구비율이 높으며 노인인구비율이 높은 지역은 젊은 경제활동인구와 가임기 여성인구비율도 상대적으로 낮기 때문에 출산율이 저조하게 나타날 수 있을 것이다.

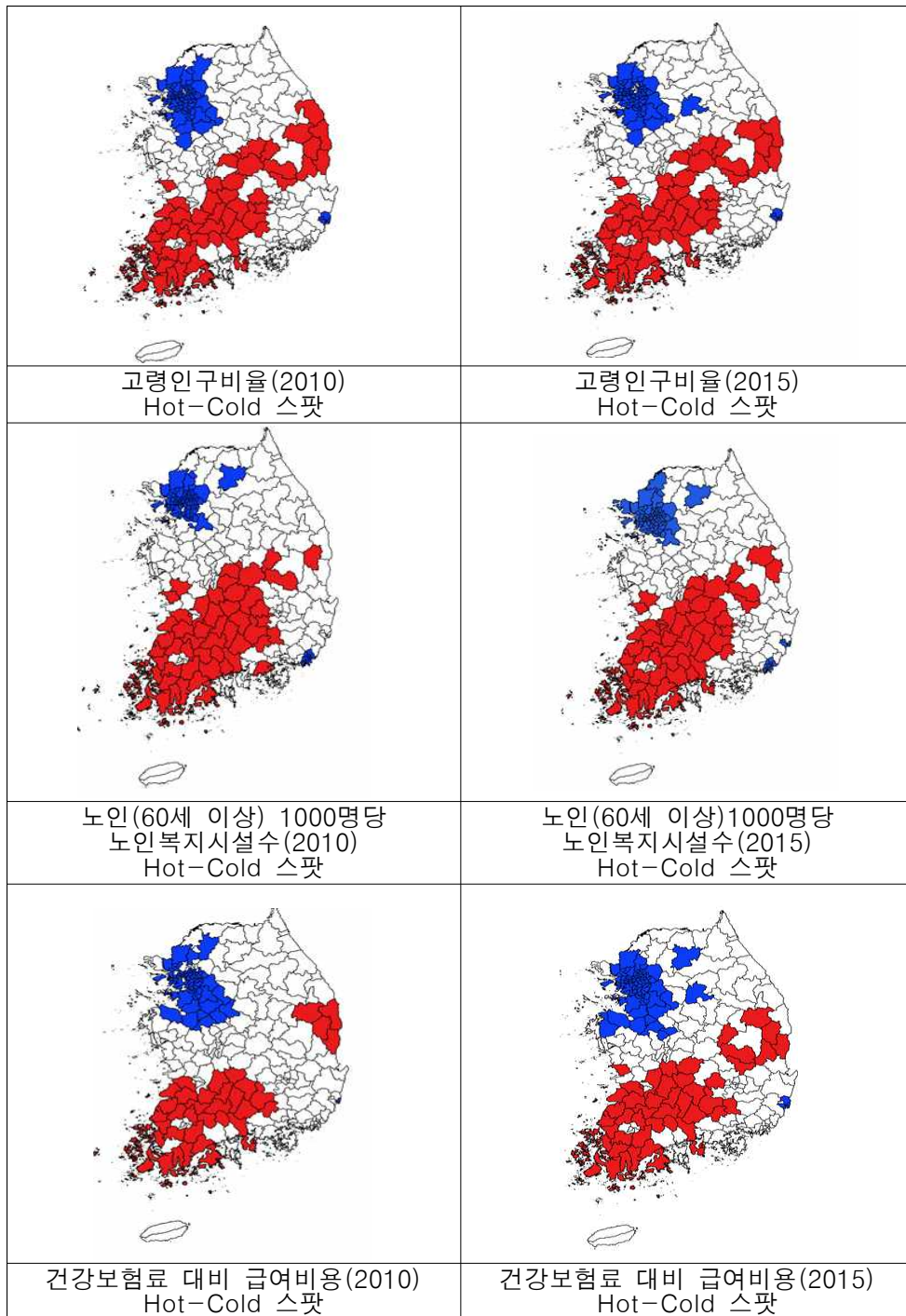


그림 10. 노인변수 Getis-Ord Gi\* 분석결과

분석 결과, 출산율에 대한 공간자기상관은 전국적 그리고 국지적으로 나타나는 것으로 확인되었으며, 상관분석을 통해 확인된 출산요인에 대한 공간분석을 실시한 결과 많은 변수들 역시 전국적 그리고 국지적으로 공간자기상관성을 나타냈다. 특히 국지적 공간분석기법인 Getis-Ord Gi\* 분석결과 모든 변수들이 연구기간에 걸쳐 전체적으로 매우 비슷한 양상을 보이는 것을 확인할 수 있었다. 이러한 추세를 분석해본다면 지역별 차별출산율과 그 요인을 설명할 수 있을 것이다. 다만, 각각의 변수에 대한 핫스팟, 콜드스팟 양상이 대도시와 지방도시를 중심으로 이루어지는 것을 볼때 지역별 출산력 분석은 공간적 이질성에 기인한 공간자기상관성을 고려한 분석이 필요하다.

## 제4절 지리가중회귀모형 분석결과

지역단위의 차별출산요인을 분석하기 위하여 앞서 제시한 바와 같이 본 연구에서는 지리가중회귀모형(GWR)을 활용하였다. 지리가중회귀모형은 공간적 이질성(도시, 지방 등)을 통제한 분석방법으로 지역별 다른 회귀분석 결과를 도출할 수 있다. 지리가중회귀모형 분석 시, 독립변수들 간의 다중공성선이 발생할 경우 결과 값이 도출되지 않기 때문에 본 연구에서는 독립변수들 간 다중공성성을 사전에 검토하여 다중공성성이 발생하지 않는 요인들로 최종 독립변수를 구성하여 분석을 실시하였다.

### 1. 연구변수 선정

다중공성선 검토는 ArcGIS를 제공하고 있는 ESRI(Environmental Systems Research Institute)에서 권고하는 방법에 따라 다중공성선 조건지수(Multicollinearity Conditional Number)가 30이하가 나오는 조합을 선택하여 다중공성선 문제를 해결하였으며, 각 인구·결혼, 여성, 지역경제·출

산환경, 노인과 관련된 연구변수를 그림 11과 같이 최종 선정하였다.

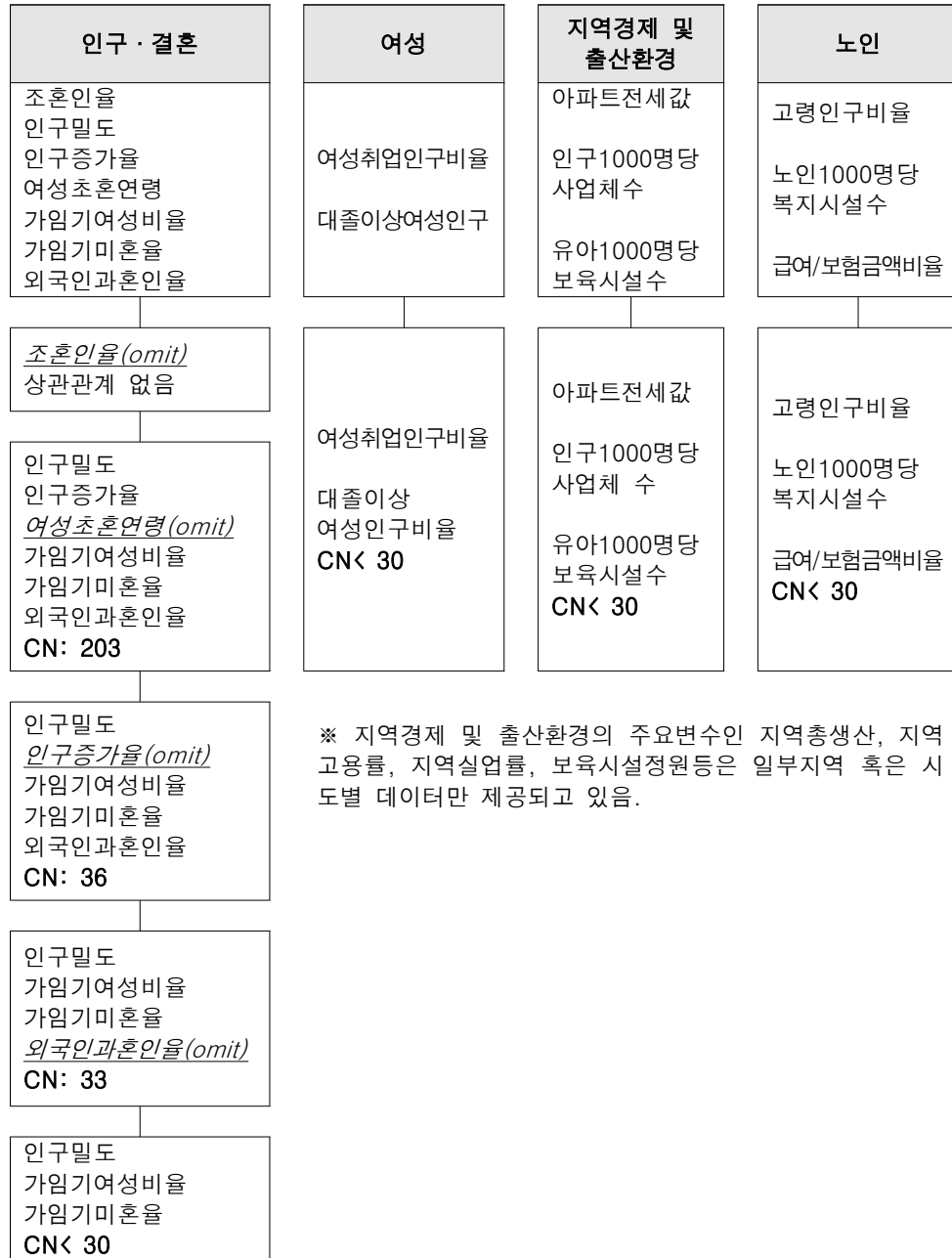


그림 11. 다중공선성 확인 및 최종 연구변수 선정

## 2. 인구·결혼변수와 지역출산율 관계 분석

인구·결혼 관련 변수는 최종적으로 인구밀도, 가임기여성비율, 가임기여성미혼비율로 선정하였다. 지역별 차별출산력 분석을 위한 지리가중회귀모형은 아래의 식과 같다.

$$\begin{aligned} Y_{\text{합계출산율}_i} = & \beta_{0i}(u) + \beta_{1i}(u)\text{인구밀도}_i \\ & + \beta_{2i}(u)\text{가임기여성비율}_i \\ & + \beta_{3i}(u)\text{가임기여성미혼율}_i \end{aligned}$$

표 12, 표 13은 지리가중회귀모형과 일반적인 OLS 모형을 사용하여 얻은 분석결과이다. 우선 OLS 모형의 추정 결과 인구밀도와 가임기 여성의 미혼비율이 높으면 출산율이 낮아지는 것으로 나타났으며, 가임기여성 인구비율의 경우 출산율에 양(+)의 효과를 주는 것으로 나타났다. 3개의 변수 모두 통계적으로 유의미한 것으로 나타났고, 2010년 자료의 분석결과 모형의 수정된  $R^2$ 는 0.50로 나타나 최종모형이 총분산의 50%를 설명할 수 있는 것으로 분석되었다. 하지만 정규성 검정인 Jarque-Bera 통계값이 유의미하게 나타났으며, 이는 잔차의 분포가 정규성을 가지지 않는다는 것을 의미한다. 또한 White-Test 통계값이 유의미하다는 것은 독립변수와 종속변수 사이에 비고정성(non-stationary)의 관계가 있음을 말해준다. 이는 일부지역에서는 특정한 독립변수가 출산율에 중요한 영향을 미치지만, 어떤 지역에서는 동일한 변수가 출산율에 미치는 영향이 약하다는 것을 의미한다. 2015년 자료의 분석결과 가임기 여성비율의 양(+) 효과가 줄어든 것을 제외하고는 2010년의 분석결과와 비슷한 결과가 도출되었다. 정규성검사와 White-Test 통계값 역시 같은 결론을 얻었으며, 결과적으로 출산율에 대한 인구·결혼 변수의 영향력이 공간적으로 다르기 때문에 지리가중회귀모형을 적용해야 함을 고려할 수 있다.

지리가중회귀모형 분석결과 지역별 출산율에 대한 인구밀도의 회귀계

수는 2010년  $-0.26 \sim -0.05$ , 2015년에는  $-0.16 \sim -0.11$ 을 보였으며 회귀계수의 표준편차 값이 작은 것으로 보아 지역별로 출산율에 대한 인구밀도 효과가 크지 않은 것으로 나타났다. 가임기여성비율과 가임기여성미혼율의 2010년 분석결과를 살펴보면 회귀계수의 표준편차 값이 0.27, 0.21로 지역별 차이가 있는 것으로 나타났으나, 2015년에는 표준편차 값이 0.01 수준으로 떨어져 가임기여성인구비율과 미혼비율로 인한 지역별 효과 차이가 줄어드는 것으로 나타났다. 그림 12는 인구·결혼 변수에 대한 지리가중회귀 분석결과를 지도화한 것이다. 그림 12는 인구·결혼 변수에 대한 지역별 회귀값을 나타내며, 색의 농도가 짙을수록 변수의 효과가 큰 것이다. 또한 왼쪽편의 그림은 2010년 자료를 지도화한 것이고, 오른쪽 그림은 2015년 자료를 지도화한 것이다. 인구밀도의 경우 전역에서 출산율에 대하여 음(-)의 효과를 보이고 있으며 2010년과 2015년에 그 효과가 전라남도지역에서 강하게 나타나고 상대적으로 동쪽지역에서 효과가 약하게 나타나는 것으로 보이지만 표준편차의 값이 적기 때문에 이로 인한 지역별 차별출산율에 미치는 영향은 다소 적었을 것으로 파악할 수 있다. 가임기여성비율은 연구지역 전역에서 출산율을 높이는 효과를 주는 것으로 나타났으며 2015년에는 부산지역에서 효과가 가장 강하게 나타났고, 서울과 경기지역에서 효과가 적은 것으로 나타났으나 지역별 표준편차가 0.01로 지역편차가 크지 않다. 하지만 2010년의 경우 가임기여성비율의 회귀계수가  $0.75 \sim 1.95$ 로 지역별 차이가 컸으며 강원도지역과 부산지역에서 효과가 강한 것으로 나타났다. 가임기여성의 미혼인구비율은 출산율과 음(-)의 관계를 보이며 2010년에는 남쪽과 북동쪽에서 회귀계수가 크게 나타났고, 중부지역에서는 음(-)이 영향이 적게 나타났다.

OLS 모형과 지리가중회귀모형의 2010년 분석결과를 비교해보면 모형의 적합성을 나타내는 AIC 값이 -60가량 줄어들었으며 모형설명력 역시 10%가량 개선된 것으로 나타났다. 2015년 분석결과 역시 GWR 모형이 OLS 모형보다 모형의 적합성과 설명력을 개선시킨 것으로 나타났다.

표 12. 2010년 인구·결혼 변수 GWR/OLS 분석결과

| 2010년         |             | GWR (N=252)                          |       |       |      | OLS (N=252) |      |         |
|---------------|-------------|--------------------------------------|-------|-------|------|-------------|------|---------|
|               |             | Min                                  | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항           |             | 1.21                                 | 1.51  | 1.92  | 0.22 | 1.56        | 0.04 | 31.61** |
| 인구밀도          |             | -0.26                                | -0.16 | -0.07 | 0.01 | -0.13       | 0.03 | -4.68** |
| 가임기<br>여성비율   |             | 0.75                                 | 1.56  | 1.95  | 0.27 | 1.42        | 0.17 | 8.35**  |
| 가임기여성<br>미혼비율 |             | -0.38                                | -0.27 | -0.06 | 0.21 | -0.36       | 0.05 | -7.07** |
| 모형<br>적합성     | AIC         | -184.28                              |       |       |      | -126.40     |      |         |
| 모형<br>설명력     | $R^2$       | 0.65                                 |       |       |      | 0.51        |      |         |
|               | $R^2_{adj}$ | 0.64                                 |       |       |      | 0.50        |      |         |
| 정규성           |             | Jarque-Bera                          |       |       |      | 310.57**    |      |         |
| 등분<br>산성      |             | White_Test                           |       |       |      | 43.276**    |      |         |
| 다중<br>공선성     |             | Multicollinearity Conditional Number |       |       |      | 22.46       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01

표 13. 2015년 인구·결혼 변수 GWR/OLS 분석결과

| 2015년         |             | GWR (N=252)                          |       |       |      | OLS (N=252) |      |         |
|---------------|-------------|--------------------------------------|-------|-------|------|-------------|------|---------|
|               |             | Min                                  | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항           |             | 1.39                                 | 1.45  | 1.56  | 0.04 | 1.48        | 0.05 | 27.11** |
| 인구밀도          |             | -0.16                                | -0.13 | -0.11 | 0.01 | -0.15       | 0.03 | -4.85** |
| 가임기<br>여성비율   |             | 0.24                                 | 0.27  | 0.35  | 0.01 | 0.3         | 0.00 | 10.43** |
| 가임기여성<br>미혼비율 |             | -0.47                                | -0.42 | -0.38 | 0.02 | -0.45       | 0.05 | -8.73** |
| 모형<br>적합성     | AIC         | -149.60                              |       |       |      | -144        |      |         |
| 모형<br>설명력     | $R^2$       | 0.58                                 |       |       |      | 0.53        |      |         |
|               | $R^2_{adj}$ | 0.61                                 |       |       |      | 0.53        |      |         |
| 정규성           |             | Jarque-Bera                          |       |       |      | 199.70**    |      |         |
| 등분<br>산성      |             | White_Test                           |       |       |      | 30.04**     |      |         |
| 다중<br>공선성     |             | Multicollinearity Conditional Number |       |       |      | 28.47       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01



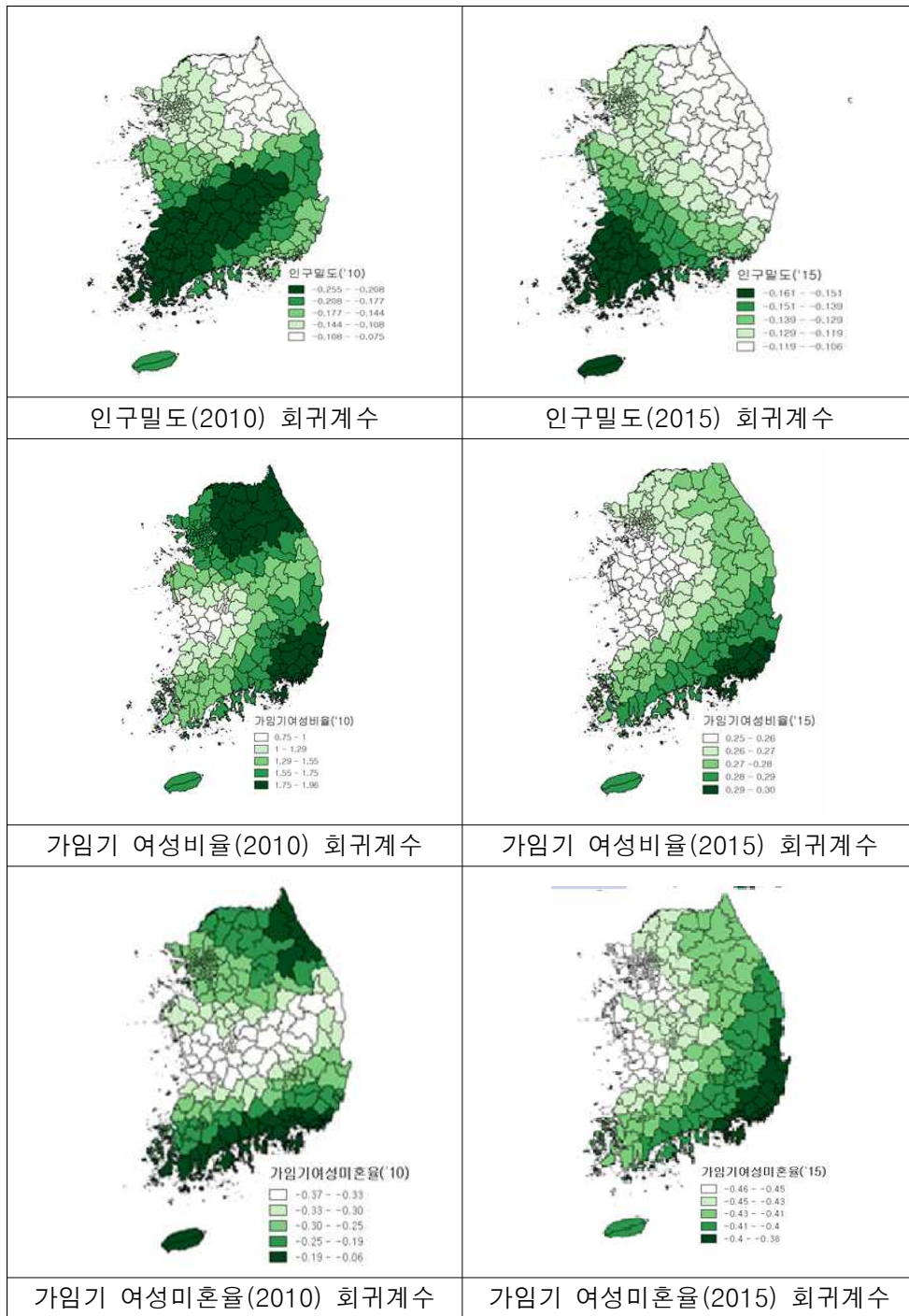


그림 12. 인구·결혼 변수 회귀계수

### 3. 여성관련 변수와 지역출산율 관계 분석

여성관련 변수는 최종적으로 여성 취업인구비율과 21-49세 여성 대학 졸업이상 인구비율을 선정하였다. 지리가중회귀분석을 위한 회귀모형은 다음과 같다.

$$Y_{\text{합계출산율}_i} = \beta_{0i}(u) + \beta_{1i}(u)\text{여성취업인구비율}_i \\ + \beta_{2i}(u)\text{여성대학졸업비율}_i$$

표 14, 표 15는 지리가중회귀모형과 일반적인 OLS 모형을 사용하여 얻은 분석결과이다. OLS 모델의 추정 결과, 2010년과 2015년 두 개의 변수가 통계적으로 유의한 것으로 나타났으며, 여성취업인구비율이 높을수록 출산율에 양(+)의 효과를 주는 반면, 여성 대학졸업인구비율이 높을수록 출산율에 음(-)의 효과를 주는 것으로 나타났다. 2010년의 경우 여성 취업인구비율이 1% 증가할 때, 출산율이 0.27만큼 증가하는 것으로 나타났다지만, 2015년에는 1% 증가가 0.63의 출산율을 높이는 것으로 나타나 여성고용이 일반적으로 출산에 긍정적인 효과를 보이는 것으로 나타났다. 반면 여성교육의 경우 2010년에 비해 2015년에는 효과가 다소 감소한 것을 확인할 수 있다. 지리가중회귀분석 결과를 살펴보면 여성취업인구비율의 표준편차가 1.74로 지역별로 여성취업으로 인한 출산율 효과 차이가 큰 것을 확인할 수 있다. 결과에서 볼 수 있듯이 OLS 분석 결과와는 달리 지역별로 여성취업인구비율이 출산율에 양(+)의 효과를 나타내는 지역이 있는 반면, 음(-)의 효과를 나타내는 지역도 있는 것을 확인할 수 있다. 그림 13을 보면 2010년 여성취업인구비율은 전라남도지역에서는 출산율을 높이는 효과를 보였지만, 강원도지역에서는 오히려 출산율을 낮추는 효과가 나타났다. 2015년에는 강원도지역과 충청도지역에서

여성취업인구비율이 출산율에 긍정적인 효과를 미쳤으나, 서울지역에서는 반대의 효과가 일어났다. 2010년과 2015년 여성취업인구비율의 지역별 회귀계수에 대한 표준편차가 크기 때문에 변수의 효과만으로 동 기간 동안 동일한 조건하에서 지역별 차별출산율이 발생한 것으로 생각해볼 수 있다. 여성대학졸업인구비율의 경우 전 지역에서 음(-)의 효과가 나타났다으며 출산율에 미치는 효과는 2010년과 2015년 유사하였으나, 2015년 표준편차가 줄어들어 여성 교육으로 인한 차별출산율 효과는 적어진 것을 확인할 수 있다. 2010년에는 교육으로 인한 음(-)의 효과가 부산에서 가장 컸고, 충청남도에서 적게 나타났다. 2015년도에는 충청남도에서는 지속적으로 여성교육으로 인한 효과가 적은 것으로 나타났지만, 강원도와 전라남도에서 교육의 효과가 크게 나타났다. OLS모형과 비교하였을 때, 지리가중회귀모형이 설명력과 적합성이 개선된 것으로 나타났다.

표 14. 2010년 여성 변수 GWR/OLS 분석결과

| 2010년          |             | GWR (N=252)                             |       |       |      | OLS (N=252) |      |         |
|----------------|-------------|---|-------|-------|------|-------------|------|---------|
|                |             | Min                                     | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항            |             | 0.89                                    | 1.49  | 2.01  | 0.23 | 1.44        | 0.16 | 8.76**  |
| 여성취업<br>인구비율   |             | -0.79                                   | 0.26  | 1.27  | 1.74 | 0.27        | 1.64 | 1.64*   |
| 대졸이상<br>여성인구비율 |             | -1.74                                   | -1.16 | -1.15 | 0.30 | -1.28       | 0.18 | -6.89** |
| 모형<br>적합성      | AIC         | -57.17                                  |       |       |      | -27.73      |      |         |
| 모형<br>설명력      | $R^2$       | 0.38                                    |       |       |      | 0.27        |      |         |
|                | $R^2_{adj}$ | 0.37                                    |       |       |      | 0.28        |      |         |
| 정규성            |             | Jarque-Bera                             |       |       |      | 53.74**     |      |         |
| 등분<br>산성       |             | White_Test                              |       |       |      | 17.23**     |      |         |
| 다중<br>공선성      |             | Multicollinearity Conditional<br>Number |       |       |      | 25.31       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01

표 15. 2015년 여성 변수 GWR/OLS 분석결과

| 2015년          |   | GWR (N=252) |       |       |      | OLS (N=252) |      |         |
|----------------|---|-------------|-------|-------|------|-------------|------|---------|
|                |   | Min         | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항            |   | -1.22       | 1.24  | 2.66  | 0.75 | 1.36        | 0.19 | 6.85**  |
| 여성취업<br>인구비율   |   | -0.97       | 0.78  | 1.33  | 1.74 | 0.63        | 0.35 | 1.78*   |
| 대졸이상<br>여성인구비율 |   | -1.76       | -1.12 | -1.02 | 0.13 | -0.98       | 0.18 | -5.51** |
| 모형<br>적합성      | AIC                                     | -1.54       |       |       |      | 1.98        |      |         |
| 모형<br>설명력      | $R^2$                                   | 0.42        |       |       |      | 0.16        |      |         |
|                | $R^2_{adj}$                             | 0.43        |       |       |      | 0.16        |      |         |
| 정규성            | Jarque-Bera                             |             |       |       |      | 74.81**     |      |         |
| 등분<br>산성       | White_Test                              |             |       |       |      | 15.79**     |      |         |
| 다중<br>공선성      | Multicollinearity Conditional<br>Number |             |       |       |      | 28.29       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01

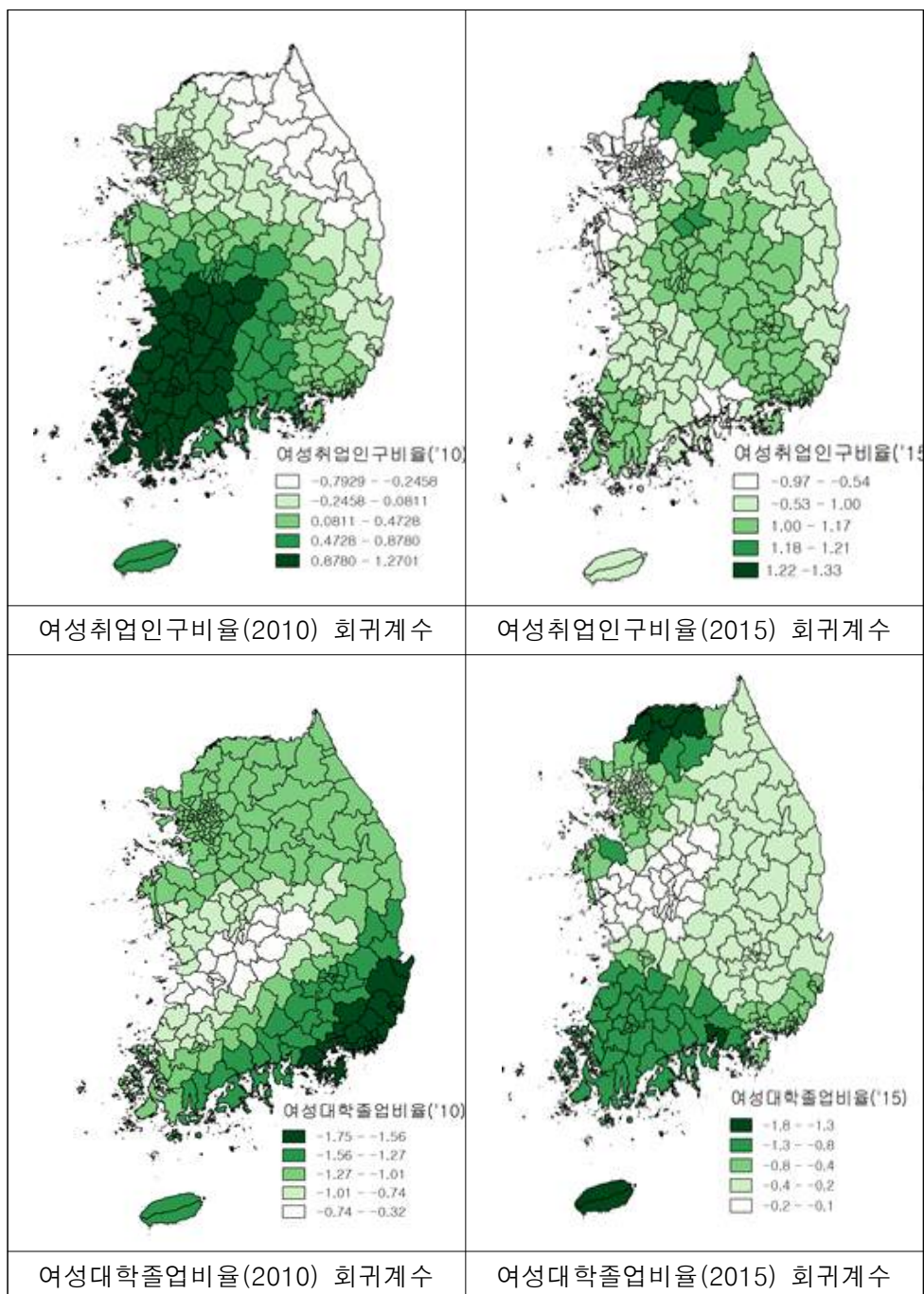


그림 13. 여성관련 변수 회귀계수

#### 4. 지역경제 및 출산환경 변수와 지역출산율 관계 분석

지역경제 및 출산환경 변수는 최종적으로 아파트 전세값과 인구 1000명당 사업체 수, 유아 1000명당 보육시설 수를 선정하여 지리가중회귀분석을 실시하였다. 분석에 활용된 회귀모형은 다음과 같다.

$$\begin{aligned} Y_{\text{합계출산율}_i} = & \beta_{0i}(u) + \beta_{1i}(u)\text{아파트전세값}(\log)_i \\ & + \beta_{2i}(u)\text{인구1000명당사업체수}(\log)_i \\ & + \beta_{3i}(u)\text{유아1000명당보육시설수}_i \end{aligned}$$

표 16, 표17은 지역경제와 출산환경에 대하여 지리가중회귀모형과 일반적인 OLS 모형을 사용하여 얻은 분석결과이다. OLS 모델의 추정 결과, 2010년에는 보육시설 수를 제외하고 아파트 전세값과 사업체 수는 출산율에 통계적으로 유의하게 음(-)의 영향을 주는 것으로 나타났다. 2015년에는 3개의 변수가 모두 통계적으로 유의하게 출산율에 음(-)의 영향을 주는 것으로 나타났다. 사업체수와 보육시설 수는 통계적으로 유의미한 결과를 보이기는 하나 출산율에 대한 효과가 매우 미비하였다. 지리가중회귀분석 결과를 살펴보면 OLS 분석결과와 마찬가지로 사업체수와 보육시설 수는 출산율에 미치는 영향력이 매우 적었으나, 지역별로 음(-)의 영향이 나타나는 곳이 있고 양(+)의 효과가 나타나는 곳이 있었다. 하지만 두 개 변수에 대한 지역별 회귀분석결과에 대한 표준편차가 아주 작았기 때문에 이로 인해 지역 간의 차별출산율의 원인으로 보기는 어려울 것으로 파악된다. 전세값의 경우, 2010년과 2015년 변수에 대한 표준편차값이 0.11, 0.12로 적은 것으로 나타났으나, 지역별로 약 -0.55 ~ 0.05의 차이가 있었기 때문에 특정지역에서는 전세값의 효과가 강하게 나타나는 것으로 나타났다. 그림 14는 지역경제와 출산환경 변수에 대한 분석결과를 지도화 한 것으로 2010년, 2015년에 공통적으로 서울과 경기

지역에서는 전세값의 음(-)의 효과가 강하게 나타나는 것으로 나타났다. 지역별 표준편차는 크지 않지만 같은 지역에서 지속적으로 전세가격의 효과가 강하게 일어난다면 전세값 효과가 같더라도 지역에 따라 산발적으로 발생하는 것보다 출산율에 대한 전세값의 효과가 장기화될 가능성이 있다. 또한 Getis-Ord Gi\* 결과에서 보았듯이 전세값의 경우 지역적 상관성이 매우 높기 때문에 전세값의 영향이 인접지역으로 퍼질 효과가 있으며, 이 경우 전세값으로 인해 지역적 차별출산력의 요인으로 작용할 가능성이 크다고 볼 수 있다. OLS 모형과 지리가중회귀모형을 비교해본 결과, 모형의 설명력과 적합성이 개선되었다.

표 16. 2010년 지역경제 및 출산환경 변수 GWR/OLS 분석결과

| 2010년             |   | GWR (N=252) |       |       |      | OLS (N=252) |      |         |
|-------------------|---|-------------|-------|-------|------|-------------|------|---------|
|                   |   | Min         | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항               |   | 1.95        | 2.30  | 2.53  | 0.17 | 2.52        | 0.13 | 19.71** |
| 아파트전세값            |   | -0.5        | -0.34 | -0.15 | 0.12 | -0.45       | 0.05 | -7.63** |
| 인구1000명당<br>사업체수  |   | -0.00       | -0.00 | -0.00 | 0.00 | -0.00       | 0.00 | -2.58** |
| 유아1000명당<br>보육시설수 |   | -0.02       | -0.00 | -0.00 | 0.00 | -0.01       | 0.00 | -2.46*  |
| 모형<br>적합성         | AIC                                     | -36.61      |       |       |      | -26         |      |         |
| 모형<br>설명력         | $R^2$                                   | 0.32        |       |       |      | 0.26        |      |         |
|                   | $R^2_{adj}$                             | 0.30        |       |       |      | 0.27        |      |         |
| 정규성               | Jarque-Bera                             |             |       |       |      | 30.58**     |      |         |
| 등분<br>산성          | White_Test                              |             |       |       |      | 33.98**     |      |         |
| 다중<br>공선성         | Multicollinearity Conditional<br>Number |             |       |       |      | 22.98       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01

표 17. 2015년 지역경제 및 출산환경 변수 GWR/OLS 분석결과

| 2015년             |   | GWR (N=252) |       |       |      | OLS (N=252) |      |         |
|-------------------|---|-------------|-------|-------|------|-------------|------|---------|
|                   |   | Min         | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항               |   | 1.76        | 2.37  | 2.94  | 0.35 | 2.47        | 0.11 | 21.61** |
| 아파트전세값            |   | -0.56       | -0.41 | -0.05 | 0.11 | -0.49       | 0.05 | -9.01** |
| 인구1000명당<br>사업체수  |   | -0.00       | -0.00 | -0.00 | 0.00 | -0.00       | 0.00 | -3.37** |
| 유아1000명당<br>보육시설수 |   | -0.03       | -0.00 | 0.00  | 0.00 | -0.00       | 0.00 | -1.55   |
| 모형<br>적합성         | AIC                                     | -69         |       |       |      | -34.93      |      |         |
| 모형<br>설명력         | $R^2$                                   | 0.44        |       |       |      | 0.29        |      |         |
|                   | $R^2_{adj}$                             | 0.40        |       |       |      | 0.28        |      |         |
| 정규성               | Jarque-Bera                             |             |       |       |      | 51.56**     |      |         |
| 등분<br>산성          | White_Test                              |             |       |       |      | 37.85**     |      |         |
| 다중<br>공선성         | Multicollinearity Conditional<br>Number |             |       |       |      | 20.86       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01



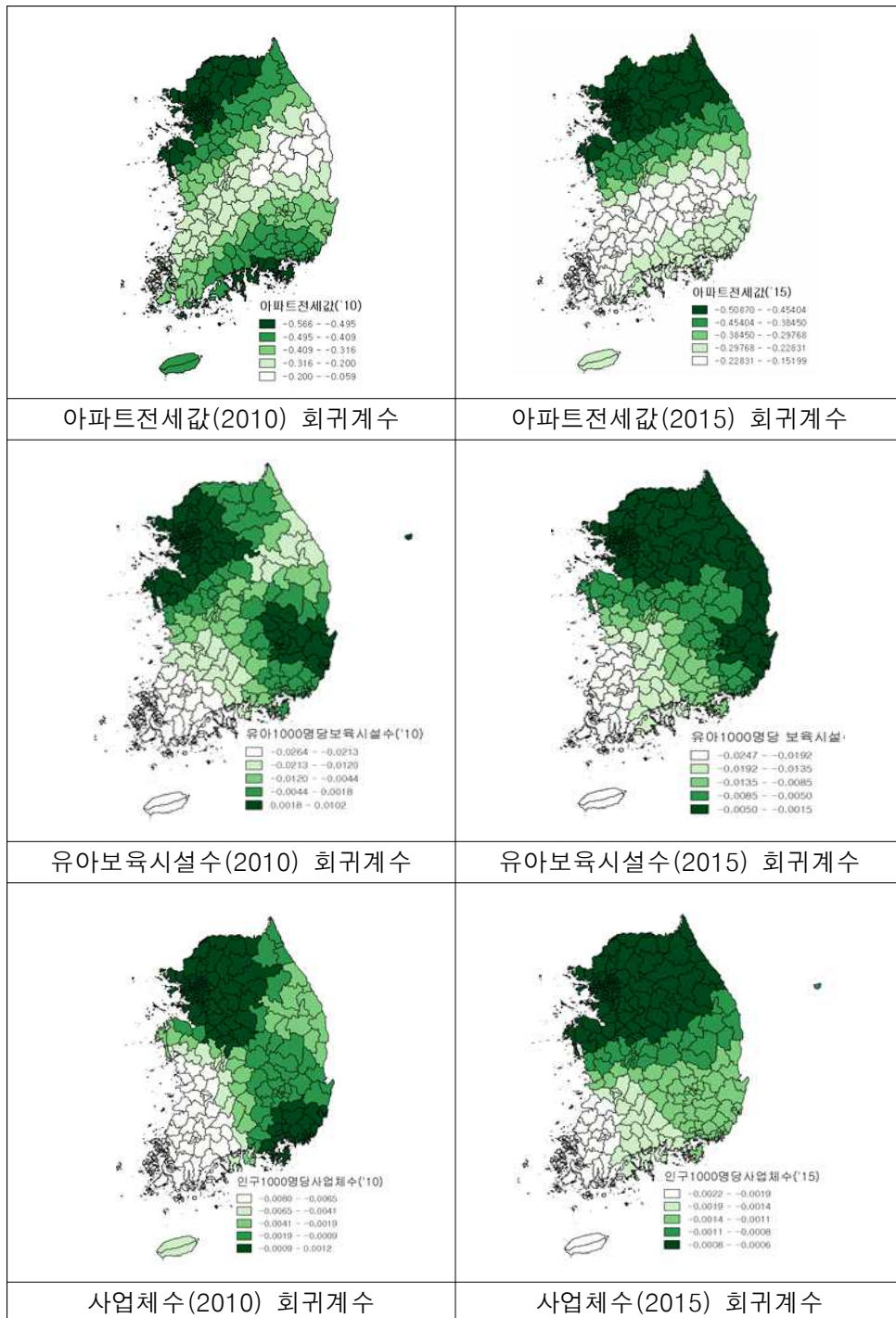


그림 14. 지역경제 및 출산환경 회귀계수

## 5. 노인관련 변수와 지역출산율 관계 분석

노인관련 변수는 다중공선성을 확인하여 최종적으로 고령인구비율과 노인 1000명당 노인복지여가시설수, 그리고 건강보험료 대비 급여지급액 비율을 선정하여 분석을 실시하였다. 분석에 활용된 회귀식은 다음과 같다.

$$\begin{aligned} Y_{\text{합계출산율}_i} = & \beta_{0i}(u) + \beta_{1i}(u)\text{고령인구비율}_i \\ & + \beta_{2i}(u)\text{노인1000명당 노인복지시설수}_i \\ & + \beta_{3i}(u)\text{급여지출액/건강보험료}_i \end{aligned}$$

표 18, 표 19는 노인 변수에 대한 지리가중회귀모형과 일반적인 OLS 모형을 사용하여 얻은 분석 결과이다. OLS 모델의 추정 결과 2010년과 2015년 모두에서 관련 변수의 효과는 0.05이하로 매우 적었으나, 2010년도 건강보험료 대비 급여지급액을 제외하고는 통계적으로 유의미한 결과가 나타났다. 동기간 동안 고령인구비율은 출산율에 음(-)의 영향을 보였고, 노인복지시설 수와 급여지출액 비율은 양(+)의 효과를 보였다. 지리가중회귀모형 분석결과를 살펴보면 지역별 회귀계수의 표준편차는 대체로 작았으며, 이는 본 연구에서 선택된 노인 변수로 인해 지역별 차별출산력을 설명하는데 한계가 있음을 설명한다. 다만, 일반적으로 고령인구비율이 높은 곳은 상대적으로 출산율 가능하게 하는 젊은 인구가 적다는 점에서 출산율이 낮을 것으로 기대할 수 있으나, 고령인구비율이 출산율에 미치는 효과만큼 노인복지시설 수가 출산율에 양(+)의 효과를 주는 것을 알 수 있다. 또한 보험료 대비 급여비가 높은 지역은 보험료 납부액 보다 급여지급액이 많은 지역으로 노인인구가 많은 곳이며 2015년도 OLS 분석결과 동 변수가 통계적으로 유의미하게 출산율에 긍정적인 효

과를 일으키는 것으로 나타났다. 그림 16을 살펴보면 2015년도에 서울지역에 노인복지시설 수가 작게나마 강하게 나타나는 것으로 보아 노인들의 건강한 활동이 양육환경 지원 등에 도움이 될 수 있음을 생각해볼 수 있다. OLS 모형과 지리가중회귀 분석결과를 비교한 결과 설명력과 모형 적합성이 개선되었으며, 고령인구비율과 건강보험료 대비 급여지급액 비율은 OLS와는 달리 지역별로 음(-)의 영향과 양(+)의 영향이 있는 곳이 있었다. 하지만 지역별 회귀계수의 표준편차가 크지 않았다.

표 18. 2010년 노인변수 GWR/OLS 분석결과

| 2010년             |                                      | GWR (N=252) |       |       |      | OLS (N=252) |      |         |
|-------------------|--------------------------------------|-------------|-------|-------|------|-------------|------|---------|
|                   |                                      | Min         | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항               |                                      | 1.39        | 1.23  | 1.47  | 0.14 | 1.21        | 0.03 | 40.95** |
| 고령인구비율            |                                      | -0.05       | -0.01 | -0.01 | 0.02 | -0.01       | 0.00 | -5.29** |
| 노인1000명당<br>복지시설수 |                                      | 0.00        | 0.03  | 0.05  | 0.01 | 0.03        | 0.00 | 12.29** |
| 건강보험료대비<br>급여지급액  |                                      | -0.10       | -0.00 | 0.32  | 0.07 | 0.01        | 0.01 | 0.65    |
| 모형<br>적합성         | AIC                                  | -134        |       |       |      | -102.01     |      |         |
| 모형<br>설명력         | $R^2$                                | 0.66        |       |       |      | 0.46        |      |         |
|                   | $R^2_{adj}$                          | 0.58        |       |       |      | 0.45        |      |         |
| 정규성               | Jarque-Bera                          |             |       |       |      | 32.56**     |      |         |
| 등분<br>산성          | White_Test                           |             |       |       |      | 31.26**     |      |         |
| 다중공<br>선성         | Multicollinearity Conditional Number |             |       |       |      | 10.08       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01

표 19. 2015년 노인변수 GWR/OLS 분석결과

| 2015년             |                                      | GWR (N=252) |       |       |      | OLS (N=252) |      |         |
|-------------------|--------------------------------------|-------------|-------|-------|------|-------------|------|---------|
|                   |                                      | Min         | Mean  | Max   | Std. | Coeff       | SdE  | t       |
| 상수항               |                                      | 0.96        | 1.32  | 1.54  | 0.15 | 1.34        | 0.03 | 37.86** |
| 고령인구비율            |                                      | -0.16       | -0.13 | -0.11 | 0.01 | -0.03       | 0.00 | -6.37** |
| 노인1000명당<br>복지시설수 |                                      | 0.2         | 0.04  | 0.07  | 0.02 | 0.03        | 0.00 | 9.94**  |
| 건강보험료대비<br>급여지급액  |                                      | -0.15       | 0.06  | 0.55  | 0.13 | 0.10        | 0.03 | 2.94*   |
| 모형<br>적합성         | AIC                                  | -126.82     |       |       |      | -70         |      |         |
| 모형<br>설명력         | $R^2$                                | 0.59        |       |       |      | 0.38        |      |         |
|                   | $R^2_{adj}$                          | 0.53        |       |       |      | 0.38        |      |         |
| 정규성               | Jarque-Bera                          |             |       |       |      | 50.83**     |      |         |
| 등분<br>산성          | White_Test                           |             |       |       |      | 31.26**     |      |         |
| 다중공<br>선성         | Multicollinearity Conditional Number |             |       |       |      | 15.37       |      |         |

※ 주 : \* p<0.05, \*\* p<0.01

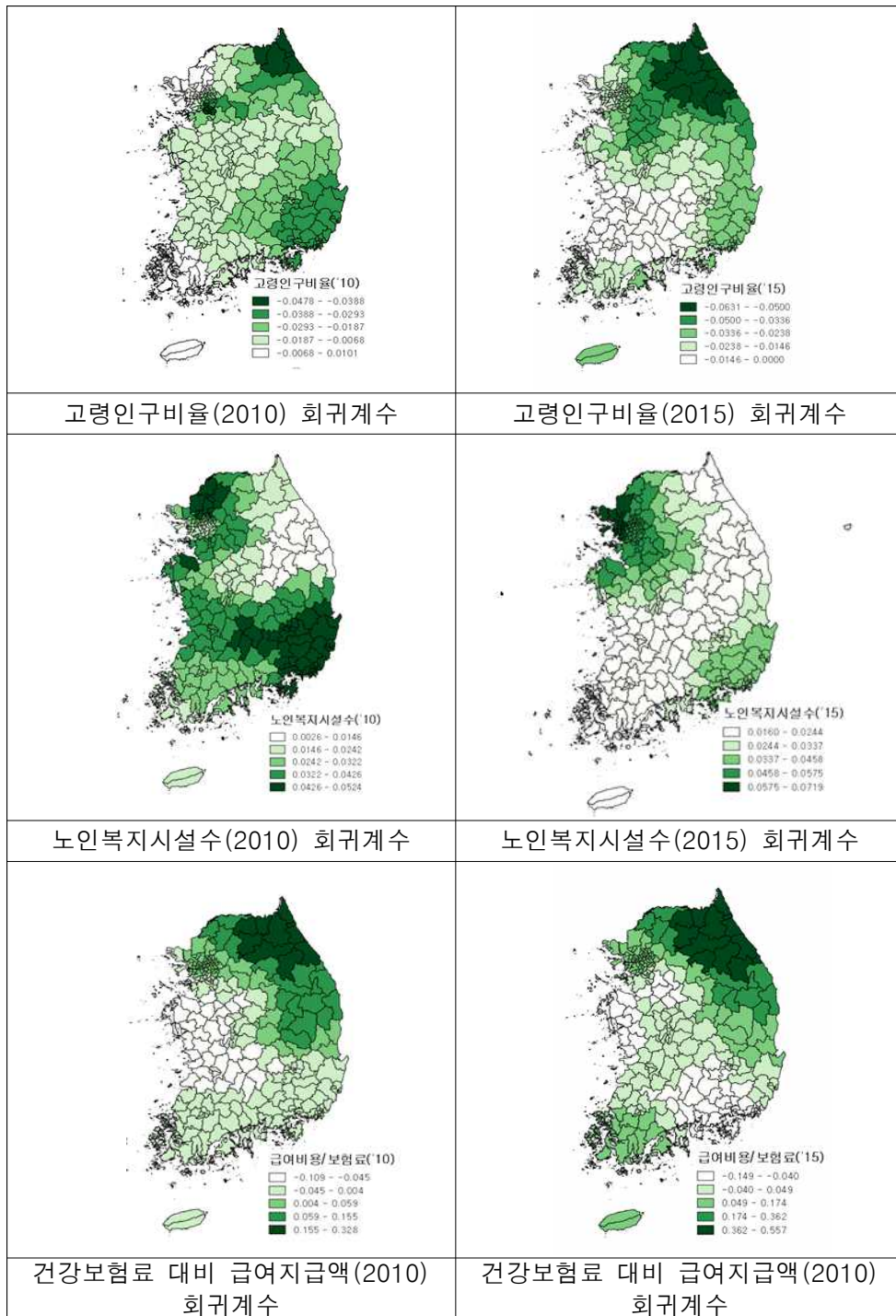


그림 15. 노인 변수 회귀계수

## 제5장 결론 및 논의

본 연구의 결론은 다음과 같다.

첫 번째 연구가설인 ‘지역별 출산율은 공간적으로 자기상관성을 갖는다.’에 대한 검증은 Moran's Index 검증을 통하여 그 결과가 0.3( $P < 0.01$ ) 이상으로 나타나 지역별 출산율이 공간자기상관성을 가지고 있음을 확인하였다. 이는 정적인(+) 공간자기상관을 나타내는 것으로 어떤 지역은 이웃한 지역과 매우 유사한 출산수준을 갖고 있다는 것을 의미한다. 또한 지역별 출산율과 함께 지역별 출산요인에 대한 공간분석을 실시한 결과, 대부분 공간자기상관성이 나타나는 것으로 확인되었다.

두 번째 연구가설인 ‘출산율이 유사한 지역끼리 군집현상을 보인다.’에 대한 검증은 Getis-Ord  $G_i^*$ 을 통해 실시하였으며, 그 결과 서울과 부산과 같은 대도시를 중심으로 출산율이 낮은 지역이 밀집한 콜드스팟이 확인되었고, 전라남도 일부에서 상대적으로 높은 출산율이 모여 있는 핫스팟 현상이 일어나는 것을 확인하였다.

마지막 연구가설인 ‘지역별 인구, 지역경제, 여성변수는 지역별 차별출산율과 유의미한 연관성을 갖는다.’에 대한 결과는 첫 번째, OLS 분석을 통해 통계적으로 유의미한 연관이 있는 것으로 확인되었다. 하지만 지리가중회귀 모형을 통해 분석을 실시한 결과, OLS 결과와 달리 여성취업인구비율의 경우 출산율에 미치는 양(+), 음(-)의 효과가 지역별로 다르게 나타나는 것이 확인되었고, 동일한 양, 음의 관계가 보이더라도 지역적 편차가 큰 변수도 확인되었다.

본 연구는 전국 252개 시군구 차별출산율을 2010년과 2015년 자료를 활용하여 분석하였다. 한국의 지역별 출산율은 대도시를 중심으로 함께

출산율이 매우 낮은 수준으로 나타났고, 전라남도과 강원도를 중심으로 상대적으로 높은 출산율을 보이는 특징이 있으며 이러한 현상은 지난 10여 년간 지속되어왔다. 이러한 현상을 지리학적인 개념으로 지리적 특징, 곧 공간적 이질성으로 인한 공간자기상관을 가진다고 말을 하며, 이는 지리적인 입지 특성에 따라 유사한 출산수준의 지역끼리 이웃해 있다는 것을 의미한다. 지리적 특징을 고려한 지역별 차별출산율을 분석하기 위하여 본 연구는 지리가중회귀모형을 활용하여 동일한 출산요인이 지역별 출산율에 미치는 영향이 다름을 확인하였다. 특히 공간자기상관으로 인해 지역별 출산요인의 영향력 수준 역시 인접한 지역끼리 유사한 값을 나타내는 것을 확인하였고, 이는 지역별 공간 효과로 인한 파급효과(Spillover Effect)가 발생하는 것을 알 수 있다. 이를 출산력 이론 중 하나인 확산효과(Diffusion Theory)와 비교해 본다면, 확산효과이론은 인접 지역간의 생각과 커뮤니케이션이 확산되어 인접지역 간의 유사한 출산력 수준을 보인다고 주장하는 반면, 본 논문의 결과는 여성의 교육수준, 아파트 전세값과 같은 정책 개입이 가능한 출산요인도 지역별 유사한 출산력 수준 확산시키는데 영향을 준다고 볼 수 있다.

한편 동일한 출산요인에 대한 영향력이 지역별로 상이한 것은, 특정 지역의 특정출산정책이 인접한 지역이 아니거나, 지역적 특성이 다른 지역에서 실시될 때 그 영향력이 같지 않을 수 있다는 것을 시사한다. 예를 들면, 최근 전라남도 일부지역의 높은 출산율에 대한 관심이 높아져 해당지역의 출산정책을 다른 지역에 반영하려는 시도는 지역별 특성에 대한 충분한 검토 이후에 이루어져야 한다.

본 연구를 통해 확인된 수도권 중심의 낮은 출산율이 밀집되어 있는 현상과 전라남도 같은 지방에 출산율이 높은 지역이 밀집되어 있는 현상은 향후에도 비슷한 양상을 보일 것으로 생각됨에 따라, 저출산 문제 해

결을 위한 지역 맞춤형 저출산 대응책이 필요하다. 가령, 낮은 출산율이 밀집된 서울은 전세값이 출산율에 미치는 영향이 크므로 출산율 하락을 줄이기 위해 주택가격 정책을 고려할 수 있다. 반면, 또 다른 낮은 출산율이 밀집된 부산에서는 전세값보다는 여성교육이 출산율에 미치는 영향이 크기 때문에 대학졸업이상의 여성 고용문제나 일-가정 양립에 대한 정책수립에 중점을 두는 등 지역의 특성과 출산요인의 효과를 고려한 저출산 대응정책을 시도해야 할 것이다.



## 제6장 연구의 한계점

본 연구의 한계점은 첫째, 본 연구는 2010년과 2015년 두 개년도의 자료를 분석하였기 때문에 출산율과 관련 변수간의 변화를 파악하는데 한계가 있었다. 향후 장기적으로 연도별 자료를 수집하여 공간분석과 시계열분석을 병합한다면 보다 흥미롭고 의미있는 연구가 이루어질 것이다. 둘째, 지리가중회귀모형은 다중공선성의 문제가 발생하면 연산이 어려워지기 때문에 이를 통제하여야 한다. 상관관계분석을 통해 출산율에 통계적으로 유의미한 상관이 있다고 확인된 출산요인 중 다중공선성이 높아 연구변수로 사용할 수 없었던 변수가 있었다. 변수들 간의 다중공선성문제를 해결할 수 있는 공간분석방법을 활용한다면 분석모형의 적합성과 설명력이 개선된 지역별 차별출산력 분석이 가능할 것이다.

위의 한계점에도 불구하고, 본 연구는 국내 출산 연구에서 부족했던 지역별 출산력에 대한 공간분석과 지역별 차별출산력에 대한 계량 지리학적(Quantitative Geography) 접근을 시도하였다는 측면에서 의의가 있다.

특히 선진국에서는 이미 GIS를 기반으로 한 연구들이 활발히 진행이 되고 있으며, 이에 대한 자료도 비교적 정교하게 구축되고 있다. 국내에서도 앞서 언급한 한계점을 보완하여 공간분석을 활용한 출산 연구들이 지속적으로 수행된다면, 지리적, 지역별 특성을 반영한 지방자치단체들의 효과적인 출산장려정책 수립과 출산환경조성에 기여할 수 있을 것이다.

## 참고 문헌

- 강창덕 (2010), “GWR 접근법을 활용한 부동산 감정평가 모형 연구: 서울시 아파트 사례로”, 「부동산연구」, 20(2): 107-132.
- 고지영, 최성일, 최인선 (2012), “경기도 시군 출산환경지수 개발 연구”, 경기도 가족여성연구원.
- 김리영, 서원석 (2016), “공간계량기법을 이용한 학령별 인구의 공간적 분포 및 지역특성 영향요인 연구”, 「한국지역개발학회지」, 28(2): 113-129.
- 김민영, 황진영 (2016), “주택가격과 출산의 시기와 수준: 우리나라 16개 시도의 실증분석”, 「보건사회연구」, 36(1): 118-142.
- 김유미, 강성홍 (2014), “지리적 가중회귀를 이용한 비만율의 지역별 변이에 관한 연구”, 「국토지리학회지」, 48(4): 425-438.
- 김혜영, 전철민 (2012), “공간구문론 및 지리적 가중회귀 기법을 이용한 지가분석”, 「한국지리정보학회지」, 15(2): 335-45.
- 류기철, 박영화 (2009), “한국여성의 출산을 변화와 출산간격 영향요인”, 「한국인구학」, 32(1): 1-23.
- 문정희, 최정락, 김하진 (2016), “출산친화적인 지역사회 환경구축 방안”, 부산여성가족개발원 연구보고서.
- 박병일 (2004), “한국의 출산력 저하 요인에 관한 연구”, 인하대학교 석사학위논문.
- 박소현, 이금숙 (2011), “사교육 시설의 수요와 공급에 나타나는 공간적 특성: 수도권 지역 사설학원을 중심으로”, 「한국경제지리학회지」, 14(1): 33-51.
- 박일수, 김은주, 홍성욱, 강성홍 (2013), “지리적 가중회귀모형을 이용한

- 지역별 심정지 발생요인에 관한 연구”, 「보건사회연구」, 33(3): 237-257.
- 박헌수 (2011), “국지가중회귀 모형을 활용한 아파트 실거래가격지수 산정방법에 관한 연구”, 「부동산연구」, 17(1): 55-66.
- 성낙일, 박선권 (2012), “우리나라 지역별 자녀 양육환경과 출산율에 관한 실증분석”, 「한국인구학」, 35(2): 73-101.
- 송용철, 박헌수 (2012), “공간계량경제 접근방법을 이용한 농지가격추정에 관한 연구: 수도권 근교 농지를 중심으로”, 「국토연구」, 72: 121-140.
- 신혜원, 최명섭, 김의준 (2012), “서울시 추가자녀 출산계획에 미치는 가구 특성의 요인분석”, 「서울도시연구」, 10(2): 33-47.
- 오운경, 강정규, 김종민 (2014), “지리가중회귀모형을 이용한 주택가격 결정요인의 지역별 특성에 관한 연구: 부산광역시를 중심으로”, 「세무회계연구」, 40: 1-17.
- 오창섭, 최성혁 (2012), “저출산 원인의 실증분석: 혼인율 하락과 초혼연령 상승의 경제적 요인”, 「복지행정논총」, 22(1): 91-125.
- 우해봉 (2012), “교육이 초혼 형성에 미치는 영향”, 「한국인구학」, 32(1): 25-50.
- 유삼현 (2007), “지역별 차별 출산력의 분석: 저출산 인과구조의 이해”, 「인구와사회」, 2(2): 53-76.
- 유삼현 (2007), “지역별 차별 출산력의 분석을 통한 저출산 인과구조의 이해”, 한양대학교대학원 석사학위논문.
- 윤정미, 최돈정 (2015), “서울시 유동인구 분포의 공간 패턴과 토지이용 특성에 관한 지리가중 회귀분석”, 「한국지형공간정보학회지」,

23(3): 77-84.

은기수 (1997), 「지역, 교육 및 직업별 차별 출산력, 한국 출산력 변천의 이해」, 서울: 일신사.

이삼식 (2016), “저출산·고령화 대책의 현황과 정책과제”, 「보건복지포럼」, 2016.1: 51-65.

이삼식, 최효진 (2012), “출산을 예측 모형 개발”, 「한국인구학」, 35(1): 77-99.

이시원, 김영기, 이성진 (2006), “우리나라 지방자치단체의 출산 감소요인에 관한 연구”, 「한국행정논집」, 18(1): 83-103.

이헌창, 승정현, 박이태, 김형중, 이경록 (2012), “저출산 고령화에 따른 지역별 인구분포와 변화요인 분석과 정책과제”, 보건사회연구원.

이희연, 심재현 (2013), 「GIS 지리정보학: 이론과 실습」, 경기도: 법문사.

정건섭 (2011), “공간계량분석방법을 이용한 부산주택특성가격 모형 비교”, 「한국비교정부학보」, 15(1): 159-184.

정영태, 강창민, 고태호, 엄상근, 임정현, 이슬, 안미영, 고윤희 (2012), 「제주특별자치도 출산영향평가 지표개발 연구」, 제주발전연구원.

정재현, 김경민 (2014), “교육의 공간 불평등 연구”, 「한국경제지리학회지」, 17(2): 385-401.

조동기 (2009), “지역 단위 조사연구와 공간정보의 활용: 지리정보시스템과 지리적 가중회귀분석을 중심으로”, 「한국조사연구학회」, 10(3): 1-19.

최돈정, 서용철 (2012), “장수의 환경생태학적 요인에 관한 지리가중회

- 귀분석” , 「한국지형공간정보학회지」 , 20(3): 57-63.
- 최은영, 박영실 (2010), “사회경제적 특성과 지역별 차별 출산력 분석” , 통계개발원연구고서.
- Boyle, P. (2003), “Population geography: does geography matter in fertility research” , *Progress in Human Geography*, 27(5): 615-626.
- Charlton, M. (2009), “Geographically Weighted Regression” , National Center for Geocomputation, National University of Ireland Maynooth, Ireland.
- Downey L. (2006). “ Environmental racial inequality in Detroit” , *Social Force* ,85:771-796
- Farber, S. and Yates. M. (2006), “A comparison of localized regression models in an hedonic price context” , *Canadian Journal of Regional Science*, 29(3), 405-420
- Fotheringham A. S. (2002). *Geographically Weighted Regression: The analysis of spatially varying relationship*, West Sussex: John Wiley&Sons Ltd.
- Getis, A. and Ord, J. (1992), “The analysis of spatial association by use of distance statistics” , *Geographical Analysis*, 24(3): 189-206.
- Ord. J. K and Getis. A (1995), “Local spatial autocorrelation statistics: Distributional issues and an application” , *Geographical Analysis*, 27(4): 286-306.
- Hank, K. (2001), “Regional fertility differences in western Germany:

- An overview of the literature and recent descriptive findings” , *International Journal of Population Geography*, 7(4): 243-257.
- Kamata, K and Iwasawa. M (2010), “Spatial variations in covariates on fertility in 2005 and 2010: Geographically Weighted Regression for small area estimates of TFR in Japan” , Paper to be presented at the XXVII IUSSP International Population Conference.
- Kulu, H., Vikat, A. and Andersson, G. (2007), “Settlement size and fertility in the Nordic countries” , *Population Studies*, 51(3): 265-285.
- Oguz. A and Melih M. P. (2006), “Geographies of a Silent Transition: A Geographically Weighted Regression Approach to Regional Fertility Differences in Turkey” , *Article in European Journal of Population*, 22(4): 399-421.
- Tobler, W. (1970), “A computer movie simulating urban growth in the detroit region” , *Economic Geography*, 46(2): 234-240.
- Wen, T. H., Chen, D. R. and Tsai, M. J. (2010), “Identifying geographical variations in poverty-obesity relationships: Empirical evidence from Taiwan” , *Geospatial Health*, 4(2): 257-265.
- Lin, C. H. and Wen, T. H. (2011), “Using Geographically Weighted Regression(GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence” , *International Journal of Environmental Research and Public Health*, 8(7): 798-815

## **Abstract**

**Myunggu Jung**

**Department of Public Health**

**Graduate School of Public Health**

**Seoul National University**

**Background:** After Total Fertility Rate dropped below to 1.3 in 2000, South Korea is experiencing long-lasting lowest low fertility phenomenon. Although South Korea government attempted to address low-fertility challenges with a series of measures aimed at boosting TFR, the record showed 1.17 in 2016, which was one of the lowest figure in the world. Interestingly, geographical fertility disparities were wider since 2000. Recent South Korea studies and research on low-fertility in South Korea are identifying the important role of regional government and suggested regional-oriented policies has to be further initiated. Therefore, this research focuses on regional fertility differences, but research was motivated from a question that areas with similar fertility level might be geographically neighbored or clustered. In other words, spatial effects or spatial auto-correlation may lead to regional fertility difference or even enlarge it.

**Objective:** The purpose of this study is adopt a quantitative geographical method in order to investigate, firstly, whether regional fertility was spatially auto-correlated with statistical significance, and, secondly, what were the determinants of regional fertility difference in 2010 and 2015.

**Method:** This study used 2010/2015 Korean Vital Statistic released from KOSIS(Korean Statistical Information Service) and 252 aggregate data of Si, Gun, Gu administrative areas were collected for analysis. TFR as a depend variable, and relevant variables were selected from 4 categories (demographic-marital characteristics, regional economy-parenting support, gender and elder)were selected as independent variables. With global regression model, such OLS (Ordinary Least Squares), spatial autocorrelation could not be handled efficiently, which often leads to unreliable inferences. The recent development of spatial statistical analysis has allowed the above methodological problems to be overcome and enabled quantitative observation of spatial effects. Hence, the Geographically Weighted Regression (GWR) model was applied to assess the heterogeneity of the relationship between regional fertility levels.

In order to investigate auto-correlation of regional fertility, Moran's Index and Getis-Ord  $G_i^*$  score (known as Hot-spot analysis) were used, where Moran's Index verifies global spatial autocorrelation and Getis-Ord  $G_i^*$  score shows local spatial autocorrelation. Then, geographically weighted regression was applied to explore determinants of regional fertility differences.

**Result:** Regional Total Fertility Rate of 252 administrative areas in 2010, 2015 showed spatial autocorrelation globally and locally. Global spatial autocorrelation of regional fertility was detected with higher than 0.3 ( $p < 0.01$ ) Moran's Index, indicating high fertility clusters near other high fertility and low fertility clusters near other low fertility. Local spatial autocorrelation of regional fertility was explored with Getis-Ord  $G_i^*$ , and the result showed that high and low fertility rate cluster spatially, in particular high fertility in metropolitan cities and



low fertility in north and south-west areas.

Geographically Weighted Regression model showed that among demographic/marital characteristics, variables related to female reproductive ages caused regional differences of fertility in 2010 but its impact faded in 2015. Of gender variables, the influence of employed female population ratio on TFR regionally and widely varied. Surprisingly, its influence showed negative in some areas and positive in particular areas. The fact that the strong impacts of Housing Rental Price on TFR were consistently detected in Seoul metropolitan areas in both 2010 and 2015 may have resulted in regionally similar patterns of fertility changes along with the effect of strong autocorrelation of Housing Rental Price. In comparison with OLS Model, the Akaike Information Criterion(AIC) decreased and adjusted R squares increased, which means the quality of statistical model improved.

**Conclusion:** This study used a quantitative geographical method to investigate regional fertility difference in 2010, 2015 South Korea. The result showed that regional fertility levels were spatially autocorrelated with statistical significance and the same independent variables had spatially different effects on fertility level, which is very likely to result in regional fertility difference in 2010/2015.

**Keywords:** Total Fertility Rate, Low fertility , Regional fertility difference, Geographically weighted regression, GIS, Quantitative geography, Spatial analysis

**Student Number:** 2015-24081